Statistical inference for 2-type doubly symmetric critical Galton–Watson processes with immigration

Márton Ispány[*], Kristóf Körmendi[⋆,⋄], Gyula Pap[⋆]

* Faculty of Informatics, University of Debrecen, Pf. 12, H-4010 Debrecen, Hungary.

⋆ Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, H-6720 Szeged, Hungary.

e-mail: ispany.marton@inf.unideb.hu (M. Ispány), kormendi@math.u-szeged.hu (K. Körmendi), papgy@math.u-szeged.hu (G. Pap).

⋄ Corresponding author.

October 19, 2012.

Abstract

In this paper the asymptotic behavior of the conditional least squares (CLS) estimators of the offspring means (α, β) and of the criticality parameter $\varrho := \alpha + \beta$ for a 2-type critical doubly symmetric positively regular Galton–Watson branching process with immigration is described.

1 Introduction

Statistical inference for critical Galton–Watson processes is available only for single-type processes, see Wei and Winnicki [16], [17] and Winnicki [18]. In the present paper the asymptotic behavior of the CLS estimators of the offspring means and criticality parameter for 2-type critical doubly symmetric positively regular Galton–Watson process with immigration is described, see Theorem 3.1. This study can be considered as the first step of examining the asymptotic behavior of the CLS estimators of parameters of multitype critical branching processes with immigration.

Let us recall the results for a single-type Galton–Watson branching process $(X_k)_{k\in\mathbb{Z}_+}$ with immigration and with initial value $X_0 = 0$. Suppose that it is critial, i.e., the offspring mean equals 1. Wei and Winnicki [16] proved a functional limit theorem $\mathcal{X}^{(n)} \xrightarrow{\mathcal{D}} \mathcal{X}$ as $n \to \infty$, where $\mathcal{X}^{(n)}_t := n^{-1} X_{\lfloor nt\rfloor}$ for $t \in \mathbb{R}_+$, $n \in \mathbb{N}$, where $\lfloor x\rfloor$ denotes the (lower) integer part of

2010 Mathematics Subject Classifications: 60J80, 62F12.

Key words and phrases: Galton–Watson branching process with immigration, conditional least squares estimator.

The authors have been supported by the Hungarian Chinese Intergovernmental S & T Cooperation Programme for 2011-2013 under Grant No. 10-1-2011-0079. K. Körmendi has been partially supported by the TÁMOP-4.2.2/B-10/1-2010-0012 project. G. Pap has been partially supported by the Hungarian Scientific Research Fund under Grant No. OTKA T-079128.

$x \in \mathbb{R}$, and $(\mathcal{X}_t)_{t\in\mathbb{R}_+}$ is a (nonnegative) diffusion process with initial value $\mathcal{X}_0 = 0$ and with generator

$$Lf(x) = m_\varepsilon f'(x) + \frac{1}{2} V_\xi x f''(x), \qquad f \in C_{\mathrm{c}}^\infty(\mathbb{R}_+),$$

where m_ε denotes the immigration mean, V_ξ denotes the offspring variance, and $C_{\mathrm{c}}^\infty(\mathbb{R}_+)$ denotes the space of infinitely differentiable functions on $\mathbb{R}_+$ with compact support. The process $(\mathcal{X}_t)_{t\in\mathbb{R}_+}$ can also be characterized as the unique strong solution of the stochastic differential equation (SDE)

$$\mathrm{d}\mathcal{X}_t = m_\varepsilon \, \mathrm{d}t + \sqrt{V_\xi \mathcal{X}_t^+} \, \mathrm{d}\mathcal{W}_t, \qquad t \in \mathbb{R}_+,$$

with initial value $\mathcal{X}_0 = 0$, where $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ is a standard Wiener process, and x^+ denotes the positive part of $x \in \mathbb{R}$. Note that this so-called square-root process is also known as Feller diffusion, or Cox–Ingersoll–Ross model in financial mathematics (see Musiela and Rutkowski [11, p. 290]). In fact, $(4V_\xi^{-1}\mathcal{X}_t)_{t\in\mathbb{R}_+}$ is the square of a $4V_\xi^{-1}m_\varepsilon$-dimensional Bessel process started at 0 (see Revuz and Yor [14, XI.1.1]).

Assuming that the immigration mean m_ε is known, for the conditional least squares estimator (CLSE)

$$\widehat{\alpha}_n(X_1, \ldots, X_n) = \frac{\sum_{k=1}^n X_{k-1}(X_k - m_\varepsilon)}{\sum_{k=1}^n X_{k-1}^2}$$

of the offspring mean based on the observations $X_1, \ldots, X_n$, one can derive

$$n(\widehat{\alpha}_n(X_1, \ldots, X_n) - 1) \xrightarrow{\mathcal{D}} \frac{\int_0^1 \mathcal{X}_t \, \mathrm{d}(\mathcal{X}_t - m_\varepsilon t)}{\int_0^1 \mathcal{X}_t^2 \, \mathrm{d}t} \qquad \text{as } n \to \infty.$$

(Wei and Winnicki [17] contains a similar result for the CLS estimator of the offspring mean when the immigration mean is unknown.)

In Section 2 we recall some preliminaries on 2-type Galton–Watson models with immigration. Section 3 contains our main results. Sections 4, 5, 6 and 7 contain the proofs. Appendix A is devoted to the CLS estimators. In Appendix B we present estimates for the moments of the processes involved. Appendix C and D is for a version of the continuous mapping theorem and for convergence of random step processes, respectively.

2 Preliminaries on 2-type Galton–Watson models with immigration

Let $\mathbb{Z}_+$, $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{R}_+$ denote the set of non-negative integers, positive integers, real numbers and non-negative real numbers, respectively. Every random variable will be defined on a fixed probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

For each $k, j \in \mathbb{Z}_+$ and $i, \ell \in \{1, 2\}$, the number of individuals of type i in the k^{th} generation will be denoted by $X_{k,i}$, the number of type ℓ offsprings produced by the j^{th}

individual who is of type i belonging to the $(k-1)^{\text{th}}$ generation will be denoted by $\xi_{k,j,i,\ell}$, and the number of type i immigrants in the k^{th} generation will be denoted by $\varepsilon_{k,i}$. Then we have

$$\begin{bmatrix} X_{k,1} \\ X_{k,2} \end{bmatrix} = \sum_{j=1}^{X_{k-1,1}} \begin{bmatrix} \xi_{k,j,1,1} \\ \xi_{k,j,1,2} \end{bmatrix} + \sum_{j=1}^{X_{k-1,2}} \begin{bmatrix} \xi_{k,j,2,1} \\ \xi_{k,j,2,2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{k,1} \\ \varepsilon_{k,2} \end{bmatrix}, \qquad k \in \mathbb{N}. \tag{2.1}$$

Here $\big\{\boldsymbol{X}_0,\, \boldsymbol{\xi}_{k,j,i},\, \boldsymbol{\varepsilon}_k : k, j \in \mathbb{N},\, i \in \{1,2\}\big\}$ are supposed to be independent, where

$$\boldsymbol{X}_k := \begin{bmatrix} X_{k,1} \\ X_{k,2} \end{bmatrix}, \qquad \boldsymbol{\xi}_{k,j,i} := \begin{bmatrix} \xi_{k,j,i,1} \\ \xi_{k,j,i,2} \end{bmatrix}, \qquad \boldsymbol{\varepsilon}_k := \begin{bmatrix} \varepsilon_{k,1} \\ \varepsilon_{k,2} \end{bmatrix}.$$

Moreover, $\{\boldsymbol{\xi}_{k,j,1} : k, j \in \mathbb{N}\}$, $\{\boldsymbol{\xi}_{k,j,2} : k, j \in \mathbb{N}\}$ and $\{\boldsymbol{\varepsilon}_k : k \in \mathbb{N}\}$ are supposed to consist of identically distributed random vectors.

We suppose $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^2) < \infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^2) < \infty$ and $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^2) < \infty$. Introduce the notations

$$\boldsymbol{m}_{\boldsymbol{\xi}_i} := \mathbb{E}\big(\boldsymbol{\xi}_{1,1,i}\big) \in \mathbb{R}_+^2, \qquad \boldsymbol{m}_{\boldsymbol{\xi}} := \begin{bmatrix} \boldsymbol{m}_{\boldsymbol{\xi}_1} & \boldsymbol{m}_{\boldsymbol{\xi}_2} \end{bmatrix} \in \mathbb{R}_+^{2\times 2}, \qquad \boldsymbol{m}_{\boldsymbol{\varepsilon}} := \mathbb{E}\big(\boldsymbol{\varepsilon}_1\big) \in \mathbb{R}_+^2,$$

$$\boldsymbol{V}_{\boldsymbol{\xi}_i} := \operatorname{Var}\big(\boldsymbol{\xi}_{1,1,i}\big) \in \mathbb{R}^{2\times 2}, \qquad \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} := \frac{1}{2}(\boldsymbol{V}_{\boldsymbol{\xi}_1} + \boldsymbol{V}_{\boldsymbol{\xi}_2}) \in \mathbb{R}^{2\times 2}, \qquad \boldsymbol{V}_{\boldsymbol{\varepsilon}} := \operatorname{Var}\big(\boldsymbol{\varepsilon}_1\big) \in \mathbb{R}^{2\times 2}.$$

Note that many authors define the offspring mean matrix as $\boldsymbol{m}_{\boldsymbol{\xi}}^\top$. For $k \in \mathbb{Z}_+$, let $\mathcal{F}_k := \sigma\big(\boldsymbol{X}_0, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_k\big)$. By (2.1),

$$\mathbb{E}(\boldsymbol{X}_k \,|\, \mathcal{F}_{k-1}) = X_{k-1,1}\, \boldsymbol{m}_{\boldsymbol{\xi}_1} + X_{k-1,2}\, \boldsymbol{m}_{\boldsymbol{\xi}_2} + \boldsymbol{m}_{\boldsymbol{\varepsilon}} = \boldsymbol{m}_{\boldsymbol{\xi}}\, \boldsymbol{X}_{k-1} + \boldsymbol{m}_{\boldsymbol{\varepsilon}}. \tag{2.2}$$

Consequently,

$$\mathbb{E}(\boldsymbol{X}_k) = \boldsymbol{m}_{\boldsymbol{\xi}}\, \mathbb{E}(\boldsymbol{X}_{k-1}) + \boldsymbol{m}_{\boldsymbol{\varepsilon}}, \qquad k \in \mathbb{N},$$

which implies

$$\mathbb{E}(\boldsymbol{X}_k) = \boldsymbol{m}_{\boldsymbol{\xi}}^k\, \mathbb{E}(\boldsymbol{X}_0) + \sum_{j=0}^{k-1} \boldsymbol{m}_{\boldsymbol{\xi}}^j\, \boldsymbol{m}_{\boldsymbol{\varepsilon}}, \qquad k \in \mathbb{N}. \tag{2.3}$$

Hence, the offspring mean matrix $\boldsymbol{m}_{\boldsymbol{\xi}}$ plays a crucial role in the asymptotic behavior of the sequence $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$. A 2-type Galton–Watson process $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ with immigration is referred to respectively as *subcritical*, *critical* or *supercritical* if $\varrho < 1$, $\varrho = 1$ or $\varrho > 1$, where ϱ denotes the spectral radius of the offspring mean matrix $\boldsymbol{m}_{\boldsymbol{\xi}}$ (see, e.g., Athreya and Ney [1, V.3] or Quine [13]). We will consider doubly symmetric 2-type Galton–Watson processes with immigration, when the offspring mean matrix has the form

$$\boldsymbol{m}_{\boldsymbol{\xi}} := \begin{bmatrix} \alpha & \beta \\ \beta & \alpha \end{bmatrix}. \tag{2.4}$$

Its spectral radius is $\varrho = \alpha + \beta$, which will be called *criticality parameter*. We will focus only on *positively regular* doubly symmetric 2-type Galton–Watson processes with immigration, i.e., when there is a positive integer $k \in \mathbb{N}$ such that the entries of $\boldsymbol{m}_{\boldsymbol{\xi}}^k$ are positive (see Kesten and Stigum [10]), which is equivalent to $\alpha > 0$ and $\beta > 0$.

For the sake of simplicity, we consider a zero start Galton–Watson process with immigration, that is, we suppose $\boldsymbol{X}_0 = \boldsymbol{0}$. The general case of nonzero initial values may be handled in a similar way, but we renounce to consider it. In the sequel we always assume $\boldsymbol{m}_\varepsilon \neq \boldsymbol{0}$, otherwise $\boldsymbol{X}_k = \boldsymbol{0}$ for all $k \in \mathbb{N}$.

3 Main results

We will use the notations

$$\boldsymbol{1} := \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in \mathbb{R}^2, \qquad \widetilde{\boldsymbol{u}} := \begin{bmatrix} 1 \\ -1 \end{bmatrix} \in \mathbb{R}^2.$$

For each $n \in \mathbb{N}$, any CLS estimator $\widehat{\varrho}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$ of the criticality parameter ϱ based on a sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ has the form

$$\widehat{\varrho}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n) = \frac{\sum_{k=1}^n \langle \boldsymbol{1}, \boldsymbol{X}_k - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \langle \boldsymbol{1}, \boldsymbol{X}_{k-1} \rangle}{\sum_{k=1}^n \langle \boldsymbol{1}, \boldsymbol{X}_{k-1} \rangle^2}, \tag{3.1}$$

whenever the sample belongs to the set

$$H_n := \left\{ (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \in (\mathbb{R}^2)^n : \sum_{k=1}^n \langle \boldsymbol{1}, \boldsymbol{x}_{k-1} \rangle^2 > 0 \right\}, \tag{3.2}$$

where $\boldsymbol{x}_0 := \boldsymbol{0}$ is the zero vector in $\mathbb{R}^2$, see Lemma A.1.

Moreover, for each $n \in \mathbb{N}$, any CLS estimator $(\widehat{\alpha}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n), \widehat{\beta}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n))$ of the offspring means (α, β) based on a sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ has the form

$$\begin{bmatrix} \widehat{\alpha}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n) \\ \widehat{\beta}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \widehat{\varrho}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n) \\ \widehat{\delta}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n) \end{bmatrix}, \tag{3.3}$$

whenever the sample belongs to the set $H_n \cap \widetilde{H}_n$, where

$$\widetilde{H}_n := \left\{ (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \in (\mathbb{R}^2)^n : \sum_{k=1}^n \langle \widetilde{\boldsymbol{u}}, \boldsymbol{x}_{k-1} \rangle^2 > 0 \right\}, \tag{3.4}$$

and

$$\widehat{\delta}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n) := \frac{\sum_{k=1}^n \langle \widetilde{\boldsymbol{u}}, \boldsymbol{X}_k - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \langle \widetilde{\boldsymbol{u}}, \boldsymbol{X}_{k-1} \rangle}{\sum_{k=1}^n \langle \widetilde{\boldsymbol{u}}, \boldsymbol{X}_{k-1} \rangle^2}, \tag{3.5}$$

see Lemma A.1.

In what follows, we always assume that $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ is a 2-type doubly symmetric Galton–Watson process with offspring means $(\alpha,\beta)\in(0,1)^2$ such that $\alpha+\beta=1$ (hence it is critical and positively regular), $\boldsymbol{X}_0=\boldsymbol{0}$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^8)<\infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^8)<\infty$, $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^8)<\infty$, and $\boldsymbol{m}_{\boldsymbol{\varepsilon}}\neq\boldsymbol{0}$. Then $\lim_{n\to\infty}\mathbb{P}((\boldsymbol{X}_1,\dots,\boldsymbol{X}_n)\in H_n)=1$. If $\langle\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\widetilde{\boldsymbol{u}},\widetilde{\boldsymbol{u}}\rangle>0$, or if $\langle\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\widetilde{\boldsymbol{u}},\widetilde{\boldsymbol{u}}\rangle=0$ and $\mathbb{E}(\langle\widetilde{\boldsymbol{u}},\boldsymbol{\varepsilon}_1\rangle^2)>0$, then $\lim_{n\to\infty}\mathbb{P}((\boldsymbol{X}_1,\dots,\boldsymbol{X}_n)\in\widetilde{H}_n)=1$, see Proposition A.4.

Let $(\mathcal{Y}_t)_{t\in\mathbb{R}_+}$ be the unique strong solution of the stochastic differential equation (SDE)

$$\mathrm{d}\mathcal{Y}_t=\langle\boldsymbol{1},\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle\,\mathrm{d}t+\sqrt{\langle\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\,\boldsymbol{1},\boldsymbol{1}\rangle\,\mathcal{Y}_t^+}\,\mathrm{d}\mathcal{W}_t,\qquad t\in\mathbb{R}_+\qquad \mathcal{Y}_0=0, \tag{3.6}$$

where $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ is a standard Wiener process.

3.1 Theorem. *We have*

$$n\big(\widehat{\varrho}_n(\boldsymbol{X}_1,\dots,\boldsymbol{X}_n)-1\big)\xrightarrow{\mathcal{D}}\frac{\int_0^1\mathcal{Y}_t\,\mathrm{d}(\mathcal{Y}_t-\langle\boldsymbol{1},\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle\,t)}{\int_0^1\mathcal{Y}_t^2\,\mathrm{d}t}\qquad \text{as } n\to\infty. \tag{3.7}$$

If $\langle\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{1},\boldsymbol{1}\rangle=0$ *then*

$$n^{3/2}\big(\widehat{\varrho}_n(\boldsymbol{X}_1,\dots,\boldsymbol{X}_n)-1\big)\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,\frac{3\langle\boldsymbol{V}_{\boldsymbol{\varepsilon}}\boldsymbol{1},\boldsymbol{1}\rangle}{\langle\boldsymbol{1},\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^2}\right)\qquad \text{as } n\to\infty. \tag{3.8}$$

If $\langle\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\widetilde{\boldsymbol{u}},\widetilde{\boldsymbol{u}}\rangle>0$ *then*

$$\begin{bmatrix} n^{1/2}(\widehat{\alpha}_n(\boldsymbol{X}_1,\dots,\boldsymbol{X}_n)-\alpha)\\ n^{1/2}(\widehat{\beta}_n(\boldsymbol{X}_1,\dots,\boldsymbol{X}_n)-\beta)\end{bmatrix}\xrightarrow{\mathcal{D}}\sqrt{\alpha\beta}\,\frac{\int_0^1\mathcal{Y}_t\,\mathrm{d}\widetilde{\mathcal{W}}_t}{\int_0^1\mathcal{Y}_t\,\mathrm{d}t}\begin{bmatrix}1\\-1\end{bmatrix}\qquad \text{as } n\to\infty, \tag{3.9}$$

where $(\widetilde{\mathcal{W}}_t)_{t\in\mathbb{R}_+}$ *is a standard Wiener process, independent from* $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$.

If $\langle\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\widetilde{\boldsymbol{u}},\widetilde{\boldsymbol{u}}\rangle=0$ *and* $\mathbb{E}(\langle\widetilde{\boldsymbol{u}},\boldsymbol{\varepsilon}_1\rangle^2)>0$ *then*

$$\begin{bmatrix} n^{1/2}(\widehat{\alpha}_n(\boldsymbol{X}_1,\dots,\boldsymbol{X}_n)-\alpha)\\ n^{1/2}(\widehat{\beta}_n(\boldsymbol{X}_1,\dots,\boldsymbol{X}_n)-\beta)\end{bmatrix}\xrightarrow{\mathcal{D}}\mathcal{N}\left(0,\frac{\langle\boldsymbol{V}_{\boldsymbol{\varepsilon}}\widetilde{\boldsymbol{u}},\widetilde{\boldsymbol{u}}\rangle}{4\,\mathbb{E}(\langle\widetilde{\boldsymbol{u}},\boldsymbol{\varepsilon}_1\rangle^2)}\right)\begin{bmatrix}1\\-1\end{bmatrix}\qquad \text{as } n\to\infty. \tag{3.10}$$

3.2 Remark. If $\langle\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\widetilde{\boldsymbol{u}},\widetilde{\boldsymbol{u}}\rangle>0$ and $\langle\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{1},\boldsymbol{1}\rangle=0$ then in (3.9) we have

$$\sqrt{\alpha\beta}\,\frac{\int_0^1\mathcal{Y}_t\,\mathrm{d}\widetilde{\mathcal{W}}_t}{\int_0^1\mathcal{Y}_t\,\mathrm{d}t}\begin{bmatrix}1\\-1\end{bmatrix}\stackrel{\mathcal{D}}{=}\mathcal{N}\left(0,\frac{4}{3}\alpha\beta\right)\begin{bmatrix}1\\-1\end{bmatrix}.$$

3.3 Remark. Note that the assumption $\langle\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{1},\boldsymbol{1}\rangle=0$ is fulfilled if and only if $\xi_{1,1,1,1}+\xi_{1,1,1,2}\stackrel{\text{a.s.}}{=}1$ and $\xi_{1,1,2,1}+\xi_{1,1,2,2}\stackrel{\text{a.s.}}{=}1$, i.e., the total number of offsprings produced by an individual of type 1 is 1, and the same holds for individuals of type 2. Indeed, $\langle\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{1},\boldsymbol{1}\rangle=\big(\langle\boldsymbol{V}_{\boldsymbol{\xi}_1}\boldsymbol{1},\boldsymbol{1}\rangle+\langle\boldsymbol{V}_{\boldsymbol{\xi}_2}\boldsymbol{1},\boldsymbol{1}\rangle\big)/2=0$ is fulfilled if and only if $\langle\boldsymbol{V}_{\boldsymbol{\xi}_1}\boldsymbol{1},\boldsymbol{1}\rangle=0$ and $\langle\boldsymbol{V}_{\boldsymbol{\xi}_2}\boldsymbol{1},\boldsymbol{1}\rangle=0$, where $\langle\boldsymbol{V}_{\boldsymbol{\xi}_i}\boldsymbol{1},\boldsymbol{1}\rangle=\mathbb{E}[\langle\boldsymbol{1},\boldsymbol{\xi}_{1,1,i}-\mathbb{E}(\boldsymbol{\xi}_{1,1,i})\rangle^2]=0$ is equivalent to $\langle\boldsymbol{1},\boldsymbol{\xi}_{1,1,i}-\mathbb{E}(\boldsymbol{\xi}_{1,1,i})\rangle=\langle\boldsymbol{1},\boldsymbol{\xi}_{1,1,i}\rangle-1\stackrel{\text{a.s.}}{=}0$ for each $i\in\{1,2\}$.

In a similar way, the assumption $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle = 0$ is fulfilled if and only if $\alpha = \beta = \frac{1}{2}$, $\xi_{1,1,1,1} \stackrel{\text{a.s.}}{=} \xi_{1,1,1,2}$ and $\xi_{1,1,2,1} \stackrel{\text{a.s.}}{=} \xi_{1,1,2,2}$, i.e., the number of offsprings of type 1 and of type 2 produced by an individual of type 1 are the same, and the same holds for individuals of type 2. Indeed, $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle = \big(\langle \boldsymbol{V}_{\boldsymbol{\xi}_1} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle + \langle \boldsymbol{V}_{\boldsymbol{\xi}_2} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle \big)/2 = 0$ is fulfilled if and only if $\langle \boldsymbol{V}_{\boldsymbol{\xi}_1} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle = 0$ and $\langle \boldsymbol{V}_{\boldsymbol{\xi}_2} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle = 0$, where $\langle \boldsymbol{V}_{\boldsymbol{\xi}_i} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle = \mathbb{E}[\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\xi}_{1,1,i} - \mathbb{E}(\boldsymbol{\xi}_{1,1,i}) \rangle^2] = 0$ is equivalent to $\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\xi}_{1,1,i} - \mathbb{E}(\boldsymbol{\xi}_{1,1,i}) \rangle = \langle \widetilde{\boldsymbol{u}}, \boldsymbol{\xi}_{1,1,i} \rangle - (\alpha - \beta) \stackrel{\text{a.s.}}{=} 0$ for each $i \in \{1, 2\}$, which imply $\alpha - \beta = 0$, since $\mathbb{P}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\xi}_{1,1,i} \rangle \in \mathbb{Z}) = 1$, $\alpha - \beta \in (-1, 1)$ by the assumptions $(\alpha, \beta) \in (0, 1)^2$ and $\alpha + \beta = 1$, and 0 is the only integer in the interval $(-1, 1)$.

Observe that the assumptions $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{1}, \boldsymbol{1} \rangle = 0$ and $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle = 0$ can not be fulfilled at the same time.

Remark that condition $\mathbb{E}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_1 \rangle^2) > 0$ fails to hold if and only if $\varepsilon_{1,1} - \varepsilon_{1,2} \stackrel{\text{a.s.}}{=} 0$, and, under the assumption $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle = 0$, this implies $X_{k,1} \stackrel{\text{a.s.}}{=} X_{k,2}$ (see Lemma A.3), when $\mathbb{P}((\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n) \in H_n \cap \widetilde{H}_n) = 0$ for all $n \in \mathbb{N}$, and hence the LSE of the offspring means (α, β) is not defined uniquely, see Appendix A. $\Box$

3.4 Remark. For each $n \in \mathbb{N}$, consider the random step process

$$\boldsymbol{\mathcal{X}}_t^{(n)} := n^{-1} \boldsymbol{X}_{\lfloor nt \rfloor}, \qquad t \in \mathbb{R}_+.$$

Theorem 5.1 implies convergence (5.3), hence

$$\boldsymbol{\mathcal{X}}^{(n)} \stackrel{\mathcal{D}}{\longrightarrow} \boldsymbol{\mathcal{X}} := \frac{1}{2} \mathcal{Y} \boldsymbol{1} \qquad \text{as } n \to \infty, \tag{3.11}$$

where the process $(\mathcal{Y}_t)_{t \in \mathbb{R}_+}$ is the unique strong solution of the SDE (3.6) with initial value $\mathcal{Y}_0 = 0$. Note that convergence (3.11) holds even if $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{1}, \boldsymbol{1} \rangle = 0$, when the unique strong solution of (3.6) is the deterministic function $\mathcal{Y}_t = \langle \boldsymbol{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \, t$, $t \in \mathbb{R}_+$. In fact, convergence (3.11) is a special case of the main result in Ispány and Pap [7, Theorem 3.1]. Indeed, the Perron vector of the offspring mean matrix $\boldsymbol{m}_{\boldsymbol{\xi}}$ is $\boldsymbol{u} = \frac{1}{2} \boldsymbol{1}$, and the unique left eigenvector $\boldsymbol{v}$ corresponding to the eigenvalue 1 of $\boldsymbol{m}_{\boldsymbol{\xi}}$ with $\boldsymbol{u}^\top \boldsymbol{v} = 1$ is $\boldsymbol{v} = \boldsymbol{1}$.

The SDE (3.6) has a unique strong solution $(\mathcal{Y}_t^{(y)})_{t \in \mathbb{R}_+}$ for all initial values $\mathcal{Y}_0^{(y)} = y \in \mathbb{R}$, and if $y \geqslant 0$, then $\mathcal{Y}_t^{(y)}$ is nonnegative for all $t \in \mathbb{R}_+$ with probability one, hence $\mathcal{Y}_t^+$ may be replaced by $\mathcal{Y}_t$ under the square root in (3.6), see, e.g., Barczy et al. [3, Remark 3.3]. $\Box$

3.5 Remark. We note that in the critical positively regular case the limit distributions for the CLS estimators of the offspring means (α, β) are concentrated on the line $\{(u, v) \in \mathbb{R}^2 : u + v = 0\}$. In order to handle the difficulty caused by this degeneracy, we use an appropriate reparametrization. Surprisingly, the scaling factor of the CLS estimators of (α, β) is always $\sqrt{n}$, which is the same as in the subcritical case. The reason of this strange phenomenon can be understood from the joint asymptotic behavior of the numerator and the denominator of the CLS estimators given in Theorems 4.1, 4.2 and 4.3. The scaling factor of the estimators of the criticality parameter ϱ is usually n, except in a particular special case of $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{1}, \boldsymbol{1} \rangle = 0$, when it is $n^{3/2}$. One of the decisive tools in deriving the needed asymptotic behavior is a good bound for the moments of the involved processes, see Corollary B.7. $\Box$

3.6 Remark. The shape of $\int_0^1 \mathcal{Y}_t \, \mathrm{d}(\mathcal{Y}_t - \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \, t) \big/ \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t$ in (3.7) is similar to the limit distribution of the Dickey–Fuller statistics for unit root test of AR(1) time series, see, e.g., Hamilton [5, 17.4.2 and 17.4.7] or Tanaka [15, (7.14) and Theorem 9.5.1]. The shape of $\int_0^1 \mathcal{Y}_t \, \mathrm{d}\widetilde{\mathcal{W}}_t \big/ \int_0^1 \mathcal{Y}_t \, \mathrm{d}t$ in (3.9) is also similar, but it contains two independent standard Wiener processes. This phenomenon is very similar to the appearance of two independent standard Wiener processes in limit theorems for CLS estimators of the variance of the offspring and immigration distributions for critical branching processes with immigration in Winnicki [18, Theorems 3.5 and 3.8]. Finally, note that the limit distribution of the CLS estimator of the criticality parameter ϱ is non-symmetric and non-normal in (3.7), and symmetric normal in (3.8), but the limit distribution of the CLS estimator of the offspring means (α, β) is always symmetric, although non-normal in (3.9). Indeed, since $(\mathcal{W}_t)_{t \in \mathbb{R}_+}$ and $(\widetilde{\mathcal{W}}_t)_{t \in \mathbb{R}_+}$ are independent, by the SDE (3.6), the processes $(\mathcal{Y}_t)_{t \in \mathbb{R}_+}$ and $(\widetilde{\mathcal{W}}_t)_{t \in \mathbb{R}_+}$ are also independent, which yields that the limit distribution of the CLS estimator of the offspring means (α, β) in (3.9) is symmetric. □

3.7 Remark. We note that an eighth order moment condition on the offspring and immigration distributions in Theorem 3.1 is supposed (i.e., we suppose $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^8) < \infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^8) < \infty$ and $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^8) < \infty$). However, it is important to remark that this condition is a technical one, we suspect that Theorem 3.1 remains true under lower order moment condition on the offspring and immigration distributions, but we renounce to consider it. □

4 Proof of the main results

Applying (2.2), let us introduce the sequence

$$\boldsymbol{M}_k := \boldsymbol{X}_k - \mathbb{E}(\boldsymbol{X}_k \,|\, \mathcal{F}_{k-1}) = \boldsymbol{X}_k - \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1} - \boldsymbol{m}_{\boldsymbol{\varepsilon}}, \qquad k \in \mathbb{N}, \tag{4.1}$$

of martingale differences with respect to the filtration $(\mathcal{F}_k)_{k \in \mathbb{Z}_+}$. By (4.1), the process $(\boldsymbol{X}_k)_{k \in \mathbb{Z}_+}$ satisfies the recursion

$$\boldsymbol{X}_k = \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1} + \boldsymbol{m}_{\boldsymbol{\varepsilon}} + \boldsymbol{M}_k, \qquad k \in \mathbb{N}. \tag{4.2}$$

Next, let us introduce the sequence

$$U_k := \langle \mathbf{1}, \boldsymbol{X}_k \rangle = X_{k,1} + X_{k,2}, \qquad k \in \mathbb{Z}_+.$$

One can observe that $U_k \geqslant 0$ for all $k \in \mathbb{Z}_+$, and

$$U_k = U_{k-1} + \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle + \langle \mathbf{1}, \boldsymbol{M}_k \rangle, \qquad k \in \mathbb{N}, \tag{4.3}$$

since $\langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1} \rangle = \mathbf{1}^\top \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1} = \mathbf{1}^\top \boldsymbol{X}_{k-1} = U_{k-1}$, because $\varrho = \alpha + \beta = 1$ implies that $\mathbf{1}$ is a left eigenvector of the mean matrix $\boldsymbol{m}_{\boldsymbol{\xi}}$ belonging to the eigenvalue 1. Hence $(U_k)_{k \in \mathbb{Z}_+}$ is a nonnegative unstable AR(1) process with positive drift $\langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle$ and with heteroscedastic innovation $(\langle \mathbf{1}, \boldsymbol{M}_k \rangle)_{k \in \mathbb{N}}$. Moreover, let

$$V_k := \langle \widetilde{\boldsymbol{u}}, \boldsymbol{X}_k \rangle = X_{k,1} - X_{k,2}, \qquad k \in \mathbb{Z}_+.$$

Note that we have

$$V_k = (\alpha-\beta)V_{k-1} + \langle \widetilde{\boldsymbol{u}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle + \langle \widetilde{\boldsymbol{u}}, \boldsymbol{M}_k\rangle, \qquad k \in \mathbb{N}, \tag{4.4}$$

since $\langle \widetilde{\boldsymbol{u}}, \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1}\rangle = \widetilde{\boldsymbol{u}}^\top \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1} = (\alpha-\beta)\widetilde{\boldsymbol{u}}^\top \boldsymbol{X}_{k-1} = (\alpha-\beta)V_{k-1}$, because $\widetilde{\boldsymbol{u}}$ is a left eigenvector of the mean matrix $\boldsymbol{m}_{\boldsymbol{\xi}}$ belonging to the eigenvalue $\alpha-\beta$. Thus $(V_k)_{k\in\mathbb{N}}$ is a stable AR(1) process with drift $\langle \widetilde{\boldsymbol{u}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle$ and with heteroscedastic innovation $(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{M}_k\rangle)_{k\in\mathbb{N}}$. Observe that

$$X_{k,1} = (U_k + V_k)/2, \qquad X_{k,2} = (U_k - V_k)/2, \qquad k \in \mathbb{Z}_+. \tag{4.5}$$

By (3.1), for each $n \in \mathbb{N}$, we have

$$\widehat{\varrho}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n) - 1 = \frac{\sum_{k=1}^n \langle \boldsymbol{1}, \boldsymbol{M}_k\rangle U_{k-1}}{\sum_{k=1}^n U_{k-1}^2},$$

whenever $(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n) \in H_n$, where H_n, $n \in \mathbb{N}$, are given in (3.2). By (3.5), for each $n \in \mathbb{N}$, we have

$$\widehat{\delta}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n) - \delta = \frac{\sum_{k=1}^n \langle \widetilde{\boldsymbol{u}}, \boldsymbol{M}_k\rangle V_{k-1}}{\sum_{k=1}^n V_{k-1}^2}, \tag{4.6}$$

whenever $(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n) \in \widetilde{H}_n$, where $\widetilde{H}_n$, $n \in \mathbb{N}$, are given in (3.4).

Theorem 3.1 will follow from the following statements by the continuous mapping theorem.

4.1 Theorem. *We have*

$$\sum_{k=1}^n \begin{bmatrix} n^{-3}U_{k-1}^2 \\ n^{-2}V_{k-1}^2 \\ n^{-2}\langle \boldsymbol{1}, \boldsymbol{M}_k\rangle U_{k-1} \\ n^{-3/2}\langle \widetilde{\boldsymbol{u}}, \boldsymbol{M}_k\rangle V_{k-1} \end{bmatrix} \stackrel{\mathcal{D}}{\longrightarrow} \begin{bmatrix} \int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t \\ \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle}{4\alpha\beta} \int_0^1 \mathcal{Y}_t \,\mathrm{d}t \\ \int_0^1 \mathcal{Y}_t \,\mathrm{d}(\mathcal{Y}_t - \langle \boldsymbol{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle\, t) \\ \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle}{2\sqrt{\alpha\beta}} \int_0^1 \mathcal{Y}_t \,\mathrm{d}\widetilde{\mathcal{W}}_t \end{bmatrix} \qquad \text{as} \quad n \to \infty.$$

4.2 Theorem. *If* $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \boldsymbol{1}, \boldsymbol{1}\rangle = 0$ *then*

$$\sum_{k=1}^n \begin{bmatrix} n^{-3}U_{k-1}^2 \\ n^{-2}V_{k-1}^2 \\ n^{-3/2}\langle \boldsymbol{1}, \boldsymbol{M}_k\rangle U_{k-1} \\ n^{-3/2}\langle \widetilde{\boldsymbol{u}}, \boldsymbol{M}_k\rangle V_{k-1} \end{bmatrix} \stackrel{\mathcal{D}}{\longrightarrow} \begin{bmatrix} \int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t \\ \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle}{4\alpha\beta} \int_0^1 \mathcal{Y}_t \,\mathrm{d}t \\ \langle \boldsymbol{V}_{\boldsymbol{\varepsilon}} \boldsymbol{1}, \boldsymbol{1}\rangle^{1/2} \int_0^1 \mathcal{Y}_t \,\mathrm{d}\widetilde{\widetilde{\mathcal{W}}}_t \\ \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle}{2\sqrt{\alpha\beta}} \int_0^1 \mathcal{Y}_t \,\mathrm{d}\widetilde{\mathcal{W}}_t \end{bmatrix} \qquad \text{as} \quad n \to \infty,$$

where $(\widetilde{\widetilde{\mathcal{W}}}_t)_{t\in\mathbb{R}_+}$ *is a standard Wiener process, independent from* $(\mathcal{W}_t)_{t\in\mathbb{R}_+}$ *and* $(\widetilde{\mathcal{W}}_t)_{t\in\mathbb{R}_+}$. *Note that* $(\mathcal{Y}_t)_{t\in\mathbb{R}_+}$ *is now the deterministic function* $\mathcal{Y}_t = \langle \boldsymbol{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle t$, $t \in \mathbb{R}_+$, *hence* $\int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t = \langle \boldsymbol{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^2/3$, $\int_0^1 \mathcal{Y}_t \,\mathrm{d}t = \langle \boldsymbol{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle/2$, $\int_0^1 \mathcal{Y}_t \,\mathrm{d}\widetilde{\widetilde{\mathcal{W}}}_t = \langle \boldsymbol{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle \int_0^1 t \,\mathrm{d}\widetilde{\widetilde{\mathcal{W}}}_t$ *and* $\int_0^1 \mathcal{Y}_t \,\mathrm{d}\widetilde{\mathcal{W}}_t = \langle \boldsymbol{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle \int_0^1 t \,\mathrm{d}\widetilde{\mathcal{W}}_t$.

4.3 Theorem. *If* $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle = 0$ *then*

$$\sum_{k=1}^{n} \begin{bmatrix} n^{-3} U_{k-1}^2 \\ n^{-1} V_{k-1}^2 \\ n^{-2} \langle \mathbf{1}, \boldsymbol{M}_k \rangle U_{k-1} \\ n^{-1/2} \langle \widetilde{\boldsymbol{u}}, \boldsymbol{M}_k \rangle V_{k-1} \end{bmatrix} \stackrel{\mathcal{D}}{\longrightarrow} \begin{bmatrix} \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t \\ \mathbb{E}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_1 \rangle^2) \\ \int_0^1 \mathcal{Y}_t \, \mathrm{d}(\mathcal{Y}_t - \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \, t) \\ [\langle \boldsymbol{V}_{\boldsymbol{\varepsilon}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle \, \mathbb{E}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_1 \rangle^2)]^{1/2} \, \widetilde{\mathcal{W}}_1 \end{bmatrix} \qquad \text{as} \quad n \to \infty.$$

5 Proof of Theorem 4.1

Consider the sequence of stochastic processes

$$\boldsymbol{\mathcal{Z}}_t^{(n)} := \begin{bmatrix} \boldsymbol{\mathcal{M}}_t^{(n)} \\ \boldsymbol{\mathcal{N}}_t^{(n)} \\ \boldsymbol{\mathcal{P}}_t^{(n)} \end{bmatrix} := \sum_{k=1}^{\lfloor nt \rfloor} \boldsymbol{Z}_k^{(n)} \qquad \text{with} \qquad \boldsymbol{Z}_k^{(n)} := \begin{bmatrix} n^{-1} \boldsymbol{M}_k \\ n^{-2} \boldsymbol{M}_k U_{k-1} \\ n^{-3/2} \boldsymbol{M}_k V_{k-1} \end{bmatrix} = \begin{bmatrix} n^{-1} \\ n^{-2} U_{k-1} \\ n^{-3/2} V_{k-1} \end{bmatrix} \otimes \boldsymbol{M}_k$$

for $t \in \mathbb{R}_+$ and $k, n \in \mathbb{N}$, where $\otimes$ denotes Kronecker product of matrices. Theorem 4.1 follows from Lemma A.2 and the following theorem (this will be explained after Theorem 5.1).

5.1 Theorem. *We have*

$$\boldsymbol{\mathcal{Z}}^{(n)} \stackrel{\mathcal{D}}{\longrightarrow} \boldsymbol{\mathcal{Z}} \qquad \text{as} \quad n \to \infty, \tag{5.1}$$

where the process $(\boldsymbol{\mathcal{Z}}_t)_{t \in \mathbb{R}_+}$ *with values in* $(\mathbb{R}^2)^3$ *is the unique strong solution of the SDE*

$$\mathrm{d}\boldsymbol{\mathcal{Z}}_t = \gamma(t, \boldsymbol{\mathcal{Z}}_t) \begin{bmatrix} \mathrm{d}\boldsymbol{\mathcal{W}}_t \\ \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \end{bmatrix}, \qquad t \in \mathbb{R}_+, \tag{5.2}$$

with initial value $\boldsymbol{\mathcal{Z}}_0 = \mathbf{0}$, *where* $(\boldsymbol{\mathcal{W}}_t)_{t \in \mathbb{R}_+}$ *and* $(\widetilde{\boldsymbol{\mathcal{W}}}_t)_{t \in \mathbb{R}_+}$ *are independent 2-dimensional standard Wiener processes, and* $\gamma : \mathbb{R}_+ \times (\mathbb{R}^2)^3 \to (\mathbb{R}^{2 \times 2})^{3 \times 2}$ *is defined by*

$$\gamma(t, \boldsymbol{x}) := \begin{bmatrix} \langle \mathbf{1}, (\boldsymbol{x}_1 + t\boldsymbol{m}_{\boldsymbol{\varepsilon}})^+ \rangle^{1/2} & 0 \\ \langle \mathbf{1}, (\boldsymbol{x}_1 + t\boldsymbol{m}_{\boldsymbol{\varepsilon}})^+ \rangle^{3/2} & 0 \\ 0 & \left(\frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle}{4\alpha\beta} \right)^{1/2} \langle \mathbf{1}, \boldsymbol{x}_1 + t\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \end{bmatrix} \otimes \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2}$$

for $t \in \mathbb{R}_+$ *and* $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) \in (\mathbb{R}^2)^3$.

(Note that the statement of Theorem 5.1 holds even if $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle = 0$, when the last 2-dimensional coordinate process of the unique strong solution $(\boldsymbol{\mathcal{Z}}_t)_{t \in \mathbb{R}_+}$ is $\mathbf{0}$.)

The SDE (5.2) has the form

$$\mathrm{d}\boldsymbol{\mathcal{Z}}_t = \begin{bmatrix} \mathrm{d}\boldsymbol{\mathcal{M}}_t \\ \mathrm{d}\boldsymbol{\mathcal{N}}_t \\ \mathrm{d}\boldsymbol{\mathcal{P}}_t \end{bmatrix} = \begin{bmatrix} \langle \mathbf{1}, (\boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\boldsymbol{\varepsilon}})^+ \rangle^{1/2} \, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \, \mathrm{d}\boldsymbol{\mathcal{W}}_t \\ \langle \mathbf{1}, (\boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\boldsymbol{\varepsilon}})^+ \rangle^{3/2} \, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \, \mathrm{d}\boldsymbol{\mathcal{W}}_t \\ \left(\frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle}{4\alpha\beta} \right)^{1/2} \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \end{bmatrix}, \qquad t \in \mathbb{R}_+.$$

Ispány and Pap [7] proved that the first 2-dimensional equation of this SDE has a unique strong solution $(\boldsymbol{\mathcal{M}}_t)_{t\in\mathbb{R}_+}$ with initial value $\boldsymbol{\mathcal{M}}_0 = \mathbf{0}$, and $(\boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\boldsymbol{\varepsilon}})^+$ may be replaced by $\boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\boldsymbol{\varepsilon}}$ (see the proof of [7, Theorem 3.1]). Thus the SDE (5.2) has a unique strong solution with initial value $\boldsymbol{\mathcal{Z}}_0 = \mathbf{0}$, and we have

$$\boldsymbol{\mathcal{Z}}_t = \begin{bmatrix} \boldsymbol{\mathcal{M}}_t \\ \boldsymbol{\mathcal{N}}_t \\ \boldsymbol{\mathcal{P}}_t \end{bmatrix} = \begin{bmatrix} \int_0^t \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^{1/2} \, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \, \mathrm{d}\boldsymbol{\mathcal{W}}_s \\ \int_0^t \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle \, \mathrm{d}\boldsymbol{\mathcal{M}}_s \\ \left(\frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle}{4\alpha\beta} \right)^{1/2} \int_0^t \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle \, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_s \end{bmatrix}, \qquad t \in \mathbb{R}_+.$$

By the method of the proof of $\mathcal{X}^{(n)} \xrightarrow{\mathcal{D}} \mathcal{X}$ in Theorem 3.1 in Barczy et al. [3], applying Lemma C.2, one can easily derive

$$\begin{bmatrix} \boldsymbol{\mathcal{X}}^{(n)} \\ \boldsymbol{\mathcal{Z}}^{(n)} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \boldsymbol{\mathcal{X}} \\ \boldsymbol{\mathcal{Z}} \end{bmatrix} \qquad \text{as } n \to \infty, \tag{5.3}$$

where

$$\boldsymbol{\mathcal{X}}_t^{(n)} := n^{-1} \boldsymbol{X}_{\lfloor nt \rfloor}, \qquad \boldsymbol{\mathcal{X}}_t := \frac{1}{2} \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \mathbf{1}, \qquad t \in \mathbb{R}_+, \qquad n \in \mathbb{N}.$$

More precisely, using that

$$\boldsymbol{X}_k = \sum_{j=1}^k \boldsymbol{m}_{\boldsymbol{\xi}}^{k-j} (\boldsymbol{M}_j + \boldsymbol{m}_{\boldsymbol{\varepsilon}}), \qquad k \in \mathbb{N},$$

we have

$$\begin{bmatrix} \boldsymbol{\mathcal{X}}^{(n)} \\ \boldsymbol{\mathcal{Z}}^{(n)} \end{bmatrix} = \psi_n(\boldsymbol{\mathcal{Z}}^{(n)}), \qquad n \in \mathbb{N},$$

where the mapping $\psi_n : \mathsf{D}(\mathbb{R}_+, (\mathbb{R}^2)^3) \to \mathsf{D}(\mathbb{R}_+, (\mathbb{R}^2)^4)$ is given by

$$\psi_n(f_1, f_2, f_3)(t) := \begin{bmatrix} \sum_{j=1}^{\lfloor nt \rfloor} \boldsymbol{m}_{\boldsymbol{\xi}}^{\lfloor nt \rfloor - j} \left(f_1\left(\frac{j}{n}\right) - f_1\left(\frac{j-1}{n}\right) + \frac{\boldsymbol{m}_{\boldsymbol{\varepsilon}}}{n} \right) \\ f_1(t) \\ f_2(t) \\ f_3(t) \end{bmatrix}$$

for $f_1, f_2, f_3 \in \mathsf{D}(\mathbb{R}_+, \mathbb{R}^2)$, $t \in \mathbb{R}_+$, $n \in \mathbb{N}$. Further, we have

$$\begin{bmatrix} \boldsymbol{\mathcal{X}} \\ \boldsymbol{\mathcal{Z}} \end{bmatrix} = \psi(\boldsymbol{\mathcal{Z}}),$$

where the mapping $\psi : \mathsf{D}(\mathbb{R}_+, (\mathbb{R}^2)^3) \to \mathsf{D}(\mathbb{R}_+, (\mathbb{R}^2)^4)$ is given by

$$\psi(f_1, f_2, f_3)(t) := \begin{bmatrix} \frac{1}{2} \langle \mathbf{1}, f_1(t) + t\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \mathbf{1} \\ f_1(t) \\ f_2(t) \\ f_3(t) \end{bmatrix}$$

for $f_1, f_2, f_3 \in \mathsf{D}(\mathbb{R}_+, \mathbb{R}^2)$ and $t \in \mathbb{R}_+$. By page 603 in Barczy et al. [3], the mappings ψ_n, $n \in \mathbb{N}$, and ψ are measurable (the latter one is continuous too), since the coordinate functions are measurable. Using page 604 in Barczy et al. [3], we obtain that the set

$$C := \big\{ f \in \mathsf{C}(\mathbb{R}_+, (\mathbb{R}^2)^3) : f(0) = \mathbf{0} \in (\mathbb{R}^2)^3 \big\}$$

has the properties $C \subseteq C_{\psi,(\psi_n)_{n\in\mathbb{N}}}$ with $C \in \mathcal{B}(\mathsf{D}(\mathbb{R}_+, (\mathbb{R}^2)^3))$ and $\mathbb{P}(\boldsymbol{\mathcal{Z}} \in C) = 1$, where $C_{\psi,(\psi_n)_{n\in\mathbb{N}}}$ is defined in Appendix C. Hence, by (5.1) and Lemma C.2, we have

$$\begin{bmatrix} \boldsymbol{\mathcal{X}}^{(n)} \\ \boldsymbol{\mathcal{Z}}^{(n)} \end{bmatrix} = \psi_n(\boldsymbol{\mathcal{Z}}^{(n)}) \xrightarrow{\mathcal{D}} \psi(\boldsymbol{\mathcal{Z}}) = \begin{bmatrix} \boldsymbol{\mathcal{X}} \\ \boldsymbol{\mathcal{Z}} \end{bmatrix} \qquad \text{as} \quad n \to \infty,$$

as desired.

Now, with the process

$$\mathcal{Y}_t := \langle \mathbf{1}, \boldsymbol{\mathcal{X}}_t \rangle = \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle, \qquad t \in \mathbb{R}_+,$$

we have

$$\boldsymbol{\mathcal{X}}_t = \frac{1}{2}\mathcal{Y}_t \mathbf{1}, \qquad t \in \mathbb{R}_+.$$

By Itô's formula we obtain that the process $(\mathcal{Y}_t)_{t\in\mathbb{R}_+}$ satisfies the SDE (3.6).

Next, similarly to the proof of (A.6), by Lemma C.3, convergence (5.3) and Lemma A.2 with $U_{k-1} = \langle \mathbf{1}, \boldsymbol{X}_{k-1} \rangle$ implies

$$\sum_{k=1}^{n} \begin{bmatrix} n^{-3} U_{k-1}^2 \\ n^{-2} V_{k-1}^2 \\ n^{-2} \langle \mathbf{1}, \boldsymbol{M}_k \rangle U_{k-1} \\ n^{-3/2} \langle \widetilde{\boldsymbol{u}}, \boldsymbol{M}_k \rangle V_{k-1} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \int_0^1 \langle \mathbf{1}, \boldsymbol{\mathcal{X}}_t \rangle^2 \, \mathrm{d}t \\ \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle}{4\alpha\beta} \int_0^1 \langle \mathbf{1}, \boldsymbol{\mathcal{X}}_t \rangle \, \mathrm{d}t \\ \int_0^1 \mathcal{Y}_t \, \mathrm{d}\langle \mathbf{1}, \boldsymbol{\mathcal{M}}_t \rangle \\ \left(\frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle}{4\alpha\beta} \right)^{1/2} \int_0^1 \mathcal{Y}_t \, \mathrm{d}\langle \widetilde{\boldsymbol{u}}, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \widetilde{\boldsymbol{\mathcal{W}}}_t \rangle \end{bmatrix} \qquad \text{as} \quad n \to \infty.$$

This limiting random vector can be written in the form as given in Theorem 4.1, since $\langle \mathbf{1}, \boldsymbol{\mathcal{X}}_t \rangle = \mathcal{Y}_t$, $\langle \mathbf{1}, \boldsymbol{\mathcal{M}}_t \rangle = \langle \mathbf{1}, \boldsymbol{\mathcal{X}}_t \rangle - \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \, t = \mathcal{Y}_t - \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \, t$ and $\langle \widetilde{\boldsymbol{u}}, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \widetilde{\boldsymbol{\mathcal{W}}}_t \rangle = \langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle^{1/2} \, \widetilde{\mathcal{W}}_t$ for all $t \in \mathbb{R}_+$ with a (one-dimensional) standard Wiener process $(\widetilde{\mathcal{W}}_t)_{t\in\mathbb{R}_+}$.

Proof of Theorem 5.1. In order to show convergence $\boldsymbol{\mathcal{Z}}^{(n)} \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{Z}}$, we apply Theorem D.1 with the special choices $\boldsymbol{\mathcal{U}} := \boldsymbol{\mathcal{Z}}$, $\boldsymbol{U}_k^{(n)} := \boldsymbol{Z}_k^{(n)}$, $n, k \in \mathbb{N}$, $(\mathcal{F}_k^{(n)})_{k\in\mathbb{Z}_+} := (\mathcal{F}_k)_{k\in\mathbb{Z}_+}$ and the function γ which is defined in Theorem 5.1. Note that the discussion after Theorem 5.1 shows that the SDE (5.2) admits a unique strong solution $(\boldsymbol{\mathcal{Z}}_t^{\boldsymbol{z}})_{t\in\mathbb{R}_+}$ for all initial values $\boldsymbol{\mathcal{Z}}_0^{\boldsymbol{z}} = \boldsymbol{z} \in (\mathbb{R}^2)^3$.

Now we show that conditions (i) and (ii) of Theorem D.1 hold. The conditional variance has the form

$$\mathbb{E}\big(\boldsymbol{Z}_k^{(n)} (\boldsymbol{Z}_k^{(n)})^\top \,|\, \mathcal{F}_{k-1}\big) = \begin{bmatrix} n^{-2} & n^{-3} U_{k-1} & n^{-5/2} V_{k-1} \\ n^{-3} U_{k-1} & n^{-4} U_{k-1}^2 & n^{-7/2} U_{k-1} V_{k-1} \\ n^{-5/2} V_{k-1} & n^{-7/2} U_{k-1} V_{k-1} & n^{-3} V_{k-1}^2 \end{bmatrix} \otimes \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \,|\, \mathcal{F}_{k-1})$$

for $n \in \mathbb{N}$, $k \in \{1, \ldots, n\}$, and $\gamma(s, \boldsymbol{\mathcal{Z}}_s^{(n)})\gamma(s, \boldsymbol{\mathcal{Z}}_s^{(n)})^\top$ has the form

$$\begin{bmatrix} \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle & \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^2 & \mathbf{0} \\ \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^2 & \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle}{4\alpha\beta} \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^2 \end{bmatrix} \otimes \overline{\boldsymbol{V}}_\xi$$

for $s \in \mathbb{R}_+$, where we used that $\langle \mathbf{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^+ = \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle$, $s \in \mathbb{R}_+$, $n \in \mathbb{N}$. Indeed, by (4.1), we get

$$\begin{aligned} \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle &= \frac{1}{n} \sum_{k=1}^{\lfloor ns \rfloor} \langle \mathbf{1}, \boldsymbol{X}_k - \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1} - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle + \langle \mathbf{1}, s\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \\ &= \frac{1}{n} \sum_{k=1}^{\lfloor ns \rfloor} \langle \mathbf{1}, \boldsymbol{X}_k - \boldsymbol{X}_{k-1} - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle + s\langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \\ &= \frac{1}{n} \langle \mathbf{1}, \boldsymbol{X}_{\lfloor ns \rfloor} \rangle + \frac{ns - \lfloor ns \rfloor}{n} \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle = \frac{1}{n} U_{\lfloor ns \rfloor} + \frac{ns - \lfloor ns \rfloor}{n} \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \end{aligned} \tag{5.4}$$

for $s \in \mathbb{R}_+$, $n \in \mathbb{N}$, since $\mathbf{1}^\top \boldsymbol{m}_{\boldsymbol{\xi}} = \mathbf{1}^\top$ implies $\langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1} \rangle = \mathbf{1}^\top \boldsymbol{m}_{\boldsymbol{\xi}} \boldsymbol{X}_{k-1} = \mathbf{1}^\top \boldsymbol{X}_{k-1} = \langle \mathbf{1}, \boldsymbol{X}_{k-1} \rangle$.

In order to check condition (i) of Theorem D.1, we need to prove that for each $T > 0$,

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^2} \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) - \int_0^t \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \, \mathrm{d}s \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{5.5}$$

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} \, \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) - \int_0^t \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^2 \, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \, \mathrm{d}s \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{5.6}$$

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^4} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \, \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) - \int_0^t \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^3 \, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \, \mathrm{d}s \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{5.7}$$

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \, \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) - \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle}{4\alpha\beta} \int_0^t \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^2 \, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \, \mathrm{d}s \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{5.8}$$

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^{5/2}} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1} \, \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{5.9}$$

$$\sup_{t \in [0,T]} \left\| \frac{1}{n^{7/2}} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1} \, \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0 \tag{5.10}$$

as $n \to \infty$.

First we show (5.5). By (5.4),

$$\int_0^t \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \, \mathrm{d}s = \frac{1}{n^2} \sum_{k=1}^{\lfloor nt \rfloor - 1} U_k + \frac{nt - \lfloor nt \rfloor}{n^2} U_{\lfloor nt \rfloor} + \frac{\lfloor nt \rfloor + (nt - \lfloor nt \rfloor)^2}{2n^2} \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle.$$

Using Lemma B.1, we obtain

$$\begin{aligned}\sum_{k=1}^{\lfloor nt\rfloor}\mathbb{E}(\boldsymbol{M}_k\boldsymbol{M}_k^\top \,|\, \mathcal{F}_{k-1}) &= \sum_{k=1}^{\lfloor nt\rfloor}\big[X_{k-1,1}\boldsymbol{V}_{\boldsymbol{\xi}_1}+X_{k-1,2}\boldsymbol{V}_{\boldsymbol{\xi}_2}+\boldsymbol{V}_{\boldsymbol{\varepsilon}}\big]\\ &= \sum_{k=1}^{\lfloor nt\rfloor}U_{k-1}\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}+\frac{1}{2}\sum_{k=1}^{\lfloor nt\rfloor}V_{k-1}(\boldsymbol{V}_{\boldsymbol{\xi}_1}-\boldsymbol{V}_{\boldsymbol{\xi}_2})+\lfloor nt\rfloor\boldsymbol{V}_{\boldsymbol{\varepsilon}}.\end{aligned}$$

Thus, in order to show (5.5), it suffices to prove

$$n^{-2}\sum_{k=1}^{\lfloor nT\rfloor}|V_k| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{5.11}$$

$$n^{-2}\sup_{t\in[0,T]}U_{\lfloor nt\rfloor} \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{5.12}$$

$$n^{-2}\sup_{t\in[0,T]}\big[\lfloor nt\rfloor+(nt-\lfloor nt\rfloor)^2\big]\to 0 \tag{5.13}$$

as $n\to\infty$. Using (B.5) with $\ell=2$, $i=0$, $j=1$ we have (5.11). Using (B.6) with $\ell=2$, $i=1$, $j=0$, we have (5.12). Clearly, (5.13) follows from $|nt-\lfloor nt\rfloor|\leqslant 1$, $n\in\mathbb{N}$, $t\in\mathbb{R}_+$, thus we conclude (5.5).

Next we turn to prove (5.6). By (5.4),

$$\begin{aligned}\int_0^t\langle\boldsymbol{1},\boldsymbol{\mathcal{M}}_s^{(n)}+s\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^2\,\mathrm{d}s &= \frac{1}{n^3}\sum_{k=1}^{\lfloor nt\rfloor-1}U_k^2+\frac{1}{n^3}\langle\boldsymbol{1},\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle\sum_{k=1}^{\lfloor nt\rfloor-1}U_k+\frac{nt-\lfloor nt\rfloor}{n^3}U_{\lfloor nt\rfloor}^2\\ &\quad+\frac{(nt-\lfloor nt\rfloor)^2}{n^3}\langle\boldsymbol{1},\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle U_{\lfloor nt\rfloor}+\frac{\lfloor nt\rfloor+(nt-\lfloor nt\rfloor)^3}{3n^3}\langle\boldsymbol{1},\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^2.\end{aligned}$$

Using Lemma B.1, we obtain

$$\begin{aligned}\sum_{k=1}^{\lfloor nt\rfloor}U_{k-1}\,\mathbb{E}(\boldsymbol{M}_k\boldsymbol{M}_k^\top \,|\, \mathcal{F}_{k-1}) &= \sum_{k=1}^{\lfloor nt\rfloor}U_{k-1}\big[X_{k-1,1}\boldsymbol{V}_{\boldsymbol{\xi}_1}+X_{k-1,2}\boldsymbol{V}_{\boldsymbol{\xi}_2}+\boldsymbol{V}_{\boldsymbol{\varepsilon}}\big]\\ &= \sum_{k=1}^{\lfloor nt\rfloor}U_{k-1}^2\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}+\frac{1}{2}\sum_{k=1}^{\lfloor nt\rfloor}U_{k-1}V_{k-1}(\boldsymbol{V}_{\boldsymbol{\xi}_1}-\boldsymbol{V}_{\boldsymbol{\xi}_2})+\sum_{k=1}^{\lfloor nt\rfloor}U_{k-1}\boldsymbol{V}_{\boldsymbol{\varepsilon}}.\end{aligned} \tag{5.14}$$

Thus, in order to show (5.6), it suffices to prove

$$n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} |U_k V_k| \xrightarrow{\mathbb{P}} 0, \tag{5.15}$$

$$n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} U_k \xrightarrow{\mathbb{P}} 0, \tag{5.16}$$

$$n^{-3/2} \sup_{t\in[0,T]} U_{\lfloor nt \rfloor} \xrightarrow{\mathbb{P}} 0, \tag{5.17}$$

$$n^{-3} \sup_{t\in[0,T]} \big[\lfloor nt \rfloor + (nt - \lfloor nt \rfloor)^3\big] \to 0 \tag{5.18}$$

as $n \to \infty$. Using (B.5) with $\ell = 2, i = 1, j = 1$ and $\ell = 2, i = 1, j = 0$, we have (5.15) and (5.16), respectively. By (B.8), we have (5.17). Clearly, (5.18) follows from $|nt - \lfloor nt \rfloor| \leqslant 1$, $n \in \mathbb{N}$, $t \in \mathbb{R}_+$, thus we conclude (5.6).

Now we turn to check (5.7). Again by (5.4), we have

$$\begin{aligned}
\int_0^t \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^3 \, \mathrm{d}s &= \frac{1}{n^4} \sum_{k=1}^{\lfloor nt \rfloor - 1} U_k^3 + \frac{3}{2n^4} \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \sum_{k=1}^{\lfloor nt \rfloor - 1} U_k^2 + \frac{1}{n^4} \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^2 \sum_{k=1}^{\lfloor nt \rfloor - 1} U_k \\
&\quad + \frac{nt - \lfloor nt \rfloor}{n^4} U_{\lfloor nt \rfloor}^3 + \frac{3(nt - \lfloor nt \rfloor)^2}{2n^4} \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle U_{\lfloor nt \rfloor}^2 \\
&\quad + \frac{(nt - \lfloor nt \rfloor)^3}{n^4} \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^2 U_{\lfloor nt \rfloor} + \frac{\lfloor nt \rfloor + (nt - \lfloor nt \rfloor)^4}{4n^4} \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^3.
\end{aligned}$$

Using Lemma B.1, we obtain

$$\begin{aligned}
&\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \, \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) = \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \big[X_{k-1,1} \boldsymbol{V}_{\boldsymbol{\xi}_1} + X_{k-1,2} \boldsymbol{V}_{\boldsymbol{\xi}_2} + \boldsymbol{V}_{\boldsymbol{\varepsilon}}\big] \\
&= \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^3 \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} + \frac{1}{2} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 V_{k-1} (\boldsymbol{V}_{\boldsymbol{\xi}_1} - \boldsymbol{V}_{\boldsymbol{\xi}_2}) + \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \boldsymbol{V}_{\boldsymbol{\varepsilon}}.
\end{aligned} \tag{5.19}$$

Thus, in order to show (5.7), it suffices to prove

$$n^{-4} \sum_{k=1}^{\lfloor nT \rfloor} |U_k^2 V_k| \xrightarrow{\mathbb{P}} 0, \tag{5.20}$$

$$n^{-4} \sum_{k=1}^{\lfloor nT \rfloor} U_k^2 \xrightarrow{\mathbb{P}} 0, \tag{5.21}$$

$$n^{-4} \sum_{k=1}^{\lfloor nT \rfloor} U_k \xrightarrow{\mathbb{P}} 0, \tag{5.22}$$

$$n^{-4/3} \sup_{t\in[0,T]} U_{\lfloor nt \rfloor} \xrightarrow{\mathbb{P}} 0, \tag{5.23}$$

$$n^{-4} \sup_{t\in[0,T]} \big[\lfloor nt \rfloor + (nt - \lfloor nt \rfloor)^4\big] \to 0 \tag{5.24}$$

as $n \to \infty$. Using (B.5) with $\ell = 4$, $i = 2$, $j = 1$, $\ell = 4$, $i = 2$, $j = 0$, and $\ell = 2$, $i = 1$, $j = 0$, we have (5.20), (5.21) and (5.22), respectively. By (B.8), we have (5.23). Clearly, (5.24) follows again from $|nt - \lfloor nt \rfloor| \leqslant 1$, $n \in \mathbb{N}$, $t \in \mathbb{R}_+$, thus we conclude (5.7).

Next we turn to prove (5.8). By (5.14), (5.15) and (5.16) we get

$$n^{-3} \sup_{t\in[0,T]} \left\| \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} \, \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) - \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \right\| \xrightarrow{\mathbb{P}} 0 \tag{5.25}$$

as $n \to \infty$ for all $T > 0$. Using (5.6), in order to prove (5.8), it is sufficient to show that

$$n^{-3} \sup_{t\in[0,T]} \left\| \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \, \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) - \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle}{4\alpha\beta} \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1}^2 \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \right\| \xrightarrow{\mathbb{P}} 0 \tag{5.26}$$

as $n \to \infty$ for all $T > 0$. As in the previous case, using Lemma B.1, we obtain

$$\begin{aligned} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \, \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) &= \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \big[X_{k-1,1} \boldsymbol{V}_{\boldsymbol{\xi}_1} + X_{k-1,2} \boldsymbol{V}_{\boldsymbol{\xi}_2} + \boldsymbol{V}_{\boldsymbol{\varepsilon}}\big] \\ &= \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1}^2 \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} + \frac{1}{2} \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^3 (\boldsymbol{V}_{\boldsymbol{\xi}_1} - \boldsymbol{V}_{\boldsymbol{\xi}_2}) + \sum_{k=1}^{\lfloor nt \rfloor} V_{k-1}^2 \boldsymbol{V}_{\boldsymbol{\varepsilon}}. \end{aligned} \tag{5.27}$$

Using (B.5) with $\ell = 6$, $i = 0$, $j = 3$ and $\ell = 4$, $i = 0$, $j = 2$, we have

$$n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} |V_k|^3 \xrightarrow{\mathbb{P}} 0, \qquad n^{-3} \sum_{k=1}^{\lfloor nT \rfloor} V_k^2 \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty,$$

hence (5.26) will follow from

$$n^{-3} \sup_{t\in[0,T]} \left\| \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}V_{k-1}^2 - \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle}{4\alpha\beta} \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2 \right\| \xrightarrow{\mathbb{P}} 0 \qquad \text{as} \quad n\to\infty \tag{5.28}$$

for all $T>0$.

The aim of the following discussion is to decompose $\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}V_{k-1}^2 - \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle}{4\alpha\beta} \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2$ as a sum of a martingale and some negligible terms. Using recursions (4.4), (4.3) and formulas (B.1) and (B.2), we obtain

$$\begin{aligned}
\mathbb{E}(U_{k-1}V_{k-1}^2 \,|\, \mathcal{F}_{k-2}) &= \mathbb{E}\Big((U_{k-2} + \langle \boldsymbol{1}, \boldsymbol{M}_{k-1} + \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle)\big((\alpha-\beta)V_{k-2} + \langle \widetilde{\boldsymbol{u}}, \boldsymbol{M}_{k-1} + \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle\big)^2 \,\Big|\, \mathcal{F}_{k-2}\Big) \\
&= (\alpha-\beta)^2 U_{k-2}V_{k-2}^2 + \widetilde{\boldsymbol{u}}^\top \mathbb{E}(\boldsymbol{M}_{k-1}\boldsymbol{M}_{k-1}^\top \,|\, \mathcal{F}_{k-2})\widetilde{\boldsymbol{u}}\, U_{k-2} \\
&\quad + \text{constant} + \text{linear combination of } \ U_{k-2}V_{k-2}, \ \ V_{k-2}^2, \ \ U_{k-2} \ \text{ and } \ V_{k-2} \\
&= (\alpha-\beta)^2 U_{k-2}V_{k-2}^2 + \langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle\, U_{k-2}^2 + \text{constant} \\
&\quad + \text{linear combination of } \ U_{k-2}V_{k-2}, \ \ V_{k-2}^2, \ \ U_{k-2} \ \text{ and } \ V_{k-2}.
\end{aligned}$$

Thus

$$\begin{aligned}
&\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}V_{k-1}^2 = \sum_{k=2}^{\lfloor nt\rfloor} \big[U_{k-1}V_{k-1}^2 - \mathbb{E}(U_{k-1}V_{k-1}^2 \,|\, \mathcal{F}_{k-2})\big] + \sum_{k=2}^{\lfloor nt\rfloor} \mathbb{E}(U_{k-1}V_{k-1}^2 \,|\, \mathcal{F}_{k-2}) \\
&= \sum_{k=2}^{\lfloor nt\rfloor} \big[U_{k-1}V_{k-1}^2 - \mathbb{E}(U_{k-1}V_{k-1}^2 \,|\, \mathcal{F}_{k-2})\big] + (\alpha-\beta)^2 \sum_{k=2}^{\lfloor nt\rfloor} U_{k-2}V_{k-2}^2 + \langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle \sum_{k=2}^{\lfloor nt\rfloor} U_{k-2}^2 \\
&\quad + \mathrm{O}(n) + \text{linear combination of } \ \sum_{k=2}^{\lfloor nt\rfloor} U_{k-2}V_{k-2}, \ \ \sum_{k=2}^{\lfloor nt\rfloor} V_{k-2}^2, \ \ \sum_{k=2}^{\lfloor nt\rfloor} U_{k-2} \ \text{ and } \ \sum_{k=2}^{\lfloor nt\rfloor} V_{k-2}.
\end{aligned}$$

Consequently,

$$\begin{aligned}
\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}V_{k-1}^2 &= \frac{1}{1-(\alpha-\beta)^2} \sum_{k=2}^{\lfloor nt\rfloor} \big[U_{k-1}V_{k-1}^2 - \mathbb{E}(U_{k-1}V_{k-1}^2 \,|\, \mathcal{F}_{k-2})\big] \\
&\quad + \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle}{1-(\alpha-\beta)^2} \sum_{k=2}^{\lfloor nt\rfloor} U_{k-2}^2 - \frac{(\alpha-\beta)^2}{1-(\alpha-\beta)^2} U_{\lfloor nt\rfloor-1}V_{\lfloor nt\rfloor-1}^2 + \mathrm{O}(n) \\
&\quad + \text{linear combination of } \ \sum_{k=2}^{\lfloor nt\rfloor} U_{k-2}V_{k-2}, \ \ \sum_{k=2}^{\lfloor nt\rfloor} V_{k-2}^2, \ \ \sum_{k=2}^{\lfloor nt\rfloor} U_{k-2} \ \text{ and } \ \sum_{k=2}^{\lfloor nt\rfloor} V_{k-2}.
\end{aligned}$$

Using (B.7) with $\ell=8$, $i=1$ and $j=2$ we have

$$n^{-3} \sup_{t\in[0,T]} \left| \sum_{k=2}^{\lfloor nt\rfloor} \big[U_{k-1}V_{k-1}^2 - \mathbb{E}(U_{k-1}V_{k-1}^2 \,|\, \mathcal{F}_{k-2})\big] \right| \xrightarrow{\mathbb{P}} 0 \qquad \text{as} \quad n\to\infty.$$

Thus, in order to show (5.28), it suffices to prove

$$n^{-3}\sum_{k=1}^{\lfloor nT\rfloor}|U_kV_k| \xrightarrow{\mathbb{P}} 0, \tag{5.29}$$

$$n^{-3}\sum_{k=1}^{\lfloor nT\rfloor}V_k^2 \xrightarrow{\mathbb{P}} 0, \tag{5.30}$$

$$n^{-3}\sum_{k=1}^{\lfloor nT\rfloor}U_k \xrightarrow{\mathbb{P}} 0, \tag{5.31}$$

$$n^{-3}\sum_{k=1}^{\lfloor nT\rfloor}|V_k| \xrightarrow{\mathbb{P}} 0, \tag{5.32}$$

$$n^{-3}\sup_{t\in[0,T]}U_{\lfloor nt\rfloor}V_{\lfloor nt\rfloor}^2 \xrightarrow{\mathbb{P}} 0, \tag{5.33}$$

$$n^{-3/2}\sup_{t\in[0,T]}U_{\lfloor nt\rfloor} \xrightarrow{\mathbb{P}} 0 \tag{5.34}$$

as $n\to\infty$. Using (B.5) with $\ell=2$, $i=1$, $j=1$; $\ell=4$, $i=0$, $j=2$; $\ell=2$, $i=1$, $j=0$, and $\ell=2$, $i=0$, $j=1$, we have (5.29), (5.30), (5.31) and (5.32). Using (B.6) with $\ell=4$, $i=1$, $j=2$ we have (5.33). By (B.8), we have (5.34). Thus we conclude (5.8).

For (5.9), consider

$$\begin{aligned}\sum_{k=1}^{\lfloor nt\rfloor}V_{k-1}\,\mathbb{E}(\boldsymbol{M}_k\boldsymbol{M}_k^\top\mid\mathcal{F}_{k-1}) &= \sum_{k=1}^{\lfloor nt\rfloor}V_{k-1}\big[X_{k-1,1}\boldsymbol{V}_{\boldsymbol{\xi}_1}+X_{k-1,2}\boldsymbol{V}_{\boldsymbol{\xi}_2}+\boldsymbol{V}_{\boldsymbol{\varepsilon}}\big]\\ &= \sum_{k=1}^{\lfloor nt\rfloor}U_{k-1}V_{k-1}\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}+\frac{1}{2}\sum_{k=1}^{\lfloor nt\rfloor}V_{k-1}^2(\boldsymbol{V}_{\boldsymbol{\xi}_1}-\boldsymbol{V}_{\boldsymbol{\xi}_2})+\sum_{k=1}^{\lfloor nt\rfloor}V_{k-1}\boldsymbol{V}_{\boldsymbol{\varepsilon}},\end{aligned} \tag{5.35}$$

where we used Lemma B.1. Using (B.5) with $\ell=4$, $i=0$, $j=2$, and $\ell=2$, $i=0$, $j=1$, we have

$$n^{-5/2}\sum_{k=1}^{\lfloor nT\rfloor}V_k^2 \xrightarrow{\mathbb{P}} 0, \qquad n^{-5/2}\sum_{k=1}^{\lfloor nT\rfloor}|V_k| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n\to\infty,$$

hence (5.9) will follow from

$$n^{-5/2}\sup_{t\in[0,T]}\left|\sum_{k=1}^{\lfloor nt\rfloor}U_{k-1}V_{k-1}\right| \xrightarrow{\mathbb{P}} 0. \tag{5.36}$$

The aim of the following discussion is to decompose $\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}V_{k-1}$ as a sum of a martingale and some negligible terms. Using the recursions (4.4), (4.3) and Lemma B.1, we obtain

$$\begin{aligned}
&\mathbb{E}(U_{k-1}V_{k-1}\,|\,\mathcal{F}_{k-2})\\
&\qquad\qquad = \mathbb{E}\Big((U_{k-2}+\langle \mathbf{1}, \boldsymbol{M}_{k-1}+\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle)\big((\alpha-\beta)V_{k-2}+\langle \widetilde{\boldsymbol{u}}, \boldsymbol{M}_{k-1}+\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle\big)\,\Big|\,\mathcal{F}_{k-2}\Big)\\
&\qquad\qquad = (\alpha-\beta)U_{k-2}V_{k-2}+\langle \widetilde{\boldsymbol{u}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle U_{k-2}+(\alpha-\beta)\langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle V_{k-2}+\mathbf{1}^\top \boldsymbol{m}_{\boldsymbol{\varepsilon}}\boldsymbol{m}_{\boldsymbol{\varepsilon}}^\top\widetilde{\boldsymbol{u}}\\
&\qquad\qquad\quad +\mathbf{1}^\top\,\mathbb{E}(\boldsymbol{M}_{k-1}\boldsymbol{M}_{k-1}^\top\,|\,\mathcal{F}_{k-2})\widetilde{\boldsymbol{u}}\\
&\qquad\qquad = (\alpha-\beta)U_{k-2}V_{k-2}+\text{constant}+\text{linear combination of } \ U_{k-2} \ \text{ and } \ V_{k-2}.
\end{aligned}$$

Thus

$$\begin{aligned}
\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}V_{k-1} &= \sum_{k=2}^{\lfloor nt\rfloor}\big[U_{k-1}V_{k-1}-\mathbb{E}(U_{k-1}V_{k-1}\,|\,\mathcal{F}_{k-2})\big]+\sum_{k=2}^{\lfloor nt\rfloor}\mathbb{E}(U_{k-1}V_{k-1}\,|\,\mathcal{F}_{k-2})\\
&= \sum_{k=2}^{\lfloor nt\rfloor}\big[U_{k-1}V_{k-1}-\mathbb{E}(U_{k-1}V_{k-1}\,|\,\mathcal{F}_{k-2})\big]+(\alpha-\beta)\sum_{k=2}^{\lfloor nt\rfloor}U_{k-2}V_{k-2}\\
&\quad +\mathrm{O}(n)+\text{linear combination of } \sum_{k=2}^{\lfloor nt\rfloor}U_{k-2} \ \text{ and } \ \sum_{k=2}^{\lfloor nt\rfloor}V_{k-2}.
\end{aligned}$$

Consequently

$$\begin{aligned}
\sum_{k=2}^{\lfloor nt\rfloor} U_{k-1}V_{k-1} &= \frac{1}{1-(\alpha-\beta)}\sum_{k=2}^{\lfloor nt\rfloor}\big[U_{k-1}V_{k-1}-\mathbb{E}(U_{k-1}V_{k-1}\,|\,\mathcal{F}_{k-2})\big]\\
&\quad -\frac{\alpha-\beta}{1-(\alpha-\beta)}U_{\lfloor nt\rfloor-1}V_{\lfloor nt\rfloor-1}+\mathrm{O}(n)+\text{linear combination of } \sum_{k=2}^{\lfloor nt\rfloor}U_{k-2} \ \text{ and } \ \sum_{k=2}^{\lfloor nt\rfloor}V_{k-2}.
\end{aligned}$$

Using (B.7) with $\ell=4$, $i=1$ and $j=1$ we have

$$n^{-5/2}\sup_{t\in[0,T]}\left|\sum_{k=2}^{\lfloor nt\rfloor}\big[U_{k-1}V_{k-1}-\mathbb{E}(U_{k-1}V_{k-1}\,|\,\mathcal{F}_{k-2})\big]\right| \stackrel{\mathbb{P}}{\longrightarrow} 0 \qquad \text{as } \ n\to\infty.$$

Thus, in order to show (5.36), it suffices to prove

$$n^{-5/2}\sum_{k=1}^{\lfloor nT\rfloor}U_k \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{5.37}$$

$$n^{-5/2}\sum_{k=1}^{\lfloor nT\rfloor}|V_k| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{5.38}$$

$$n^{-5/2}\sup_{t\in[0,T]}|U_{\lfloor nt\rfloor}V_{\lfloor nt\rfloor}| \stackrel{\mathbb{P}}{\longrightarrow} 0 \tag{5.39}$$

as $n \to \infty$. Using (B.5) with $\ell = 2$, $i = 1$, $j = 0$, and $\ell = 2$, $i = 0$, $j = 1$, we have (5.37) and (5.38). Using (B.6) with $\ell = 3$, $i = 1$, $j = 1$ we have (5.39), thus we conclude (5.9).

Convergence (5.10) can be handled in the same way as (5.9). For completeness we present all of the details. By Lemma B.1, we have

$$\begin{aligned}\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}V_{k-1}\,\mathbb{E}(\boldsymbol{M}_k^2 \,|\, \mathcal{F}_{k-1}) &= \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}V_{k-1}\big[X_{k-1,1}\boldsymbol{V}_{\boldsymbol{\xi}_1} + X_{k-1,2}\boldsymbol{V}_{\boldsymbol{\xi}_2} + \boldsymbol{V}_{\boldsymbol{\varepsilon}}\big]\\ &= \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2V_{k-1}\overline{\boldsymbol{V}}_{\boldsymbol{\xi}} + \frac{1}{2}\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}V_{k-1}^2(\boldsymbol{V}_{\boldsymbol{\xi}_1} - \boldsymbol{V}_{\boldsymbol{\xi}_2}) + \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}V_{k-1}\boldsymbol{V}_{\boldsymbol{\varepsilon}}.\end{aligned} \tag{5.40}$$

Using (B.5) with $\ell = 4$, $i = 1$, $j = 2$, and $\ell = 2$, $i = 1$, $j = 1$, we have

$$n^{-7/2}\sum_{k=1}^{\lfloor nT\rfloor} U_{k-1}V_{k-1}^2 \stackrel{\mathbb{P}}{\longrightarrow} 0, \qquad n^{-7/2}\sum_{k=1}^{\lfloor nT\rfloor} |U_{k-1}V_{k-1}| \stackrel{\mathbb{P}}{\longrightarrow} 0 \qquad \text{as } n \to \infty,$$

hence (5.10) will follow from

$$n^{-7/2}\sup_{t\in[0,T]}\left|\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2V_{k-1}\right| \stackrel{\mathbb{P}}{\longrightarrow} 0 \qquad \text{as } n \to \infty. \tag{5.41}$$

The aim of the following discussion is to decompose $\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2V_{k-1}$ as a sum of a martingale and some negligible terms. Using recursions (4.4), (4.3) and Lemma B.1, we obtain

$$\begin{aligned}\mathbb{E}(U_{k-1}^2V_{k-1} \,|\, \mathcal{F}_{k-2}) &= \mathbb{E}\Big((U_{k-2} + \langle \boldsymbol{1}, \boldsymbol{M}_{k-1} + \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle)^2\big((\alpha-\beta)V_{k-2} + \langle \widetilde{\boldsymbol{u}}, \boldsymbol{M}_{k-1} + \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle\big)\,\Big|\,\mathcal{F}_{k-2}\Big)\\ &= (\alpha-\beta)U_{k-2}^2V_{k-2} + \text{constant}\\ &\qquad + \text{linear combination of } U_{k-2},\ V_{k-2},\ U_{k-2}^2,\ V_{k-2}^2 \text{ and } U_{k-2}V_{k-2}.\end{aligned}$$

Thus

$$\begin{aligned}\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2V_{k-1} &= \sum_{k=2}^{\lfloor nt\rfloor}\big[U_{k-1}^2V_{k-1} - \mathbb{E}(U_{k-1}^2V_{k-1} \,|\, \mathcal{F}_{k-2})\big] + \sum_{k=2}^{\lfloor nt\rfloor}\mathbb{E}(U_{k-1}^2V_{k-1} \,|\, \mathcal{F}_{k-2})\\ &= \sum_{k=2}^{\lfloor nt\rfloor}\big[U_{k-1}^2V_{k-1} - \mathbb{E}(U_{k-1}^2V_{k-1} \,|\, \mathcal{F}_{k-2})\big] + (\alpha-\beta)\sum_{k=2}^{\lfloor nt\rfloor} U_{k-2}^2V_{k-2} + \mathrm{O}(n)\\ &\qquad + \text{linear combination of } \sum_{k=1}^{\lfloor nt\rfloor} U_{k-2},\ \sum_{k=1}^{\lfloor nt\rfloor} V_{k-2},\ \sum_{k=1}^{\lfloor nt\rfloor} U_{k-2}^2,\ \sum_{k=1}^{\lfloor nt\rfloor} V_{k-2}^2 \text{ and } \sum_{k=1}^{\lfloor nt\rfloor} U_{k-2}V_{k-2}.\end{aligned}$$

Consequently

$$\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2 V_{k-1} = \frac{1}{2\beta} \sum_{k=2}^{\lfloor nt\rfloor} \left[U_{k-1}^2 V_{k-1} - \mathbb{E}(U_{k-1}^2 V_{k-1} \,|\, \mathcal{F}_{k-2})\right] - \frac{\alpha-\beta}{2\beta} U_{\lfloor nt\rfloor-1}^2 V_{\lfloor nt\rfloor-1} + \mathrm{O}(n)$$

$$+ \text{linear combination of } \sum_{k=1}^{\lfloor nt\rfloor} U_{k-2}, \ \sum_{k=1}^{\lfloor nt\rfloor} V_{k-2}, \ \sum_{k=1}^{\lfloor nt\rfloor} U_{k-2}^2, \ \sum_{k=1}^{\lfloor nt\rfloor} V_{k-2}^2 \ \text{ and } \ \sum_{k=1}^{\lfloor nt\rfloor} U_{k-2} V_{k-2}.$$

Using (B.7) with $\ell = 8$, $i = 2$ and $j = 1$ we have

$$n^{-7/2} \sup_{t\in[0,T]} \left| \sum_{k=2}^{\lfloor nt\rfloor} \left[U_{k-1}^2 V_{k-1} - \mathbb{E}(U_{k-1}^2 V_{k-1} \,|\, \mathcal{F}_{k-2})\right] \right| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty.$$

Thus, in order to show (5.41), it suffices to prove

$$n^{-7/2} \sum_{k=1}^{\lfloor nT\rfloor} U_k \xrightarrow{\mathbb{P}} 0, \tag{5.42}$$

$$n^{-7/2} \sum_{k=1}^{\lfloor nT\rfloor} U_k^2 \xrightarrow{\mathbb{P}} 0, \tag{5.43}$$

$$n^{-7/2} \sum_{k=1}^{\lfloor nT\rfloor} |V_k| \xrightarrow{\mathbb{P}} 0, \tag{5.44}$$

$$n^{-7/2} \sum_{k=1}^{\lfloor nT\rfloor} V_k^2 \xrightarrow{\mathbb{P}} 0, \tag{5.45}$$

$$n^{-7/2} \sum_{k=1}^{\lfloor nT\rfloor} |U_k V_k| \xrightarrow{\mathbb{P}} 0, \tag{5.46}$$

$$n^{-7/2} \sup_{t\in[0,T]} |U_{\lfloor nt\rfloor}^2 V_{\lfloor nt\rfloor}| \xrightarrow{\mathbb{P}} 0 \tag{5.47}$$

as $n \to \infty$. Here (5.42), (5.43), (5.44), (5.45) and (5.46) follow by (B.5), and (5.47) by (B.6), thus we conclude (5.10).

Finally, we check condition (ii) of Theorem D.1, i.e., the conditional Lindeberg condition

$$\sum_{k=1}^{\lfloor nT\rfloor} \mathbb{E}\left(\|\boldsymbol{Z}_k^{(n)}\|^2 \mathbb{1}_{\{\|\boldsymbol{Z}_k^{(n)}\|>\theta\}} \,\middle|\, \mathcal{F}_{k-1}\right) \xrightarrow{\mathbb{P}} 0 \qquad \text{for all } \theta > 0 \text{ and } T > 0. \tag{5.48}$$

We have $\mathbb{E}\left(\|\boldsymbol{Z}_k^{(n)}\|^2 \mathbb{1}_{\{\|\boldsymbol{Z}_k^{(n)}\|>\theta\}} \,\middle|\, \mathcal{F}_{k-1}\right) \leqslant \theta^{-2}\, \mathbb{E}\left(\|\boldsymbol{Z}_k^{(n)}\|^4 \,\middle|\, \mathcal{F}_{k-1}\right)$ and

$$\|\boldsymbol{Z}_k^{(n)}\|^4 \leqslant 3\left(n^{-4} + n^{-8} U_{k-1}^4 + n^{-6} V_{k-1}^4\right) \|\boldsymbol{M}_{k-1}\|^4.$$

Hence

$$\sum_{k=1}^{\lfloor nT\rfloor} \mathbb{E}\left(\|\boldsymbol{Z}_k^{(n)}\|^2 \mathbb{1}_{\{\|\boldsymbol{Z}_k^{(n)}\|>\theta\}}\right) \to 0 \qquad \text{as } n\to\infty \text{ for all } \theta>0 \text{ and } T>0,$$

since $\mathbb{E}(\|\boldsymbol{M}_k\|^4) = \mathrm{O}(k^2)$, $\mathbb{E}(\|\boldsymbol{M}_k\|^4 U_{k-1}^4) \leqslant \sqrt{\mathbb{E}(\|\boldsymbol{M}_k\|^8)\,\mathbb{E}(U_{k-1}^8)} = \mathrm{O}(k^6)$ and $\mathbb{E}(\|\boldsymbol{M}_k\|^4 V_{k-1}^4) \leqslant \sqrt{\mathbb{E}(\|\boldsymbol{M}_k\|^8)\,\mathbb{E}(V_{k-1}^8)} = \mathrm{O}(k^4)$ by Corollary B.7. Here we call the attention to the fact that our eighth order moment conditions $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^8)<\infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^8)<\infty$ and $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^8)<\infty$ are used for applying Corollary B.7. This yields (5.48). □

6 Proof of Theorem 4.2

This is similar to the proof of Theorem 4.1. Consider the sequence of stochastic processes

$$\boldsymbol{\mathcal{Z}}_t^{(n)} := \begin{bmatrix} \boldsymbol{\mathcal{M}}_t^{(n)} \\ \mathcal{N}_t^{(n)} \\ \boldsymbol{\mathcal{P}}_t^{(n)} \end{bmatrix} := \sum_{k=1}^{\lfloor nt\rfloor} \boldsymbol{Z}_k^{(n)} \qquad \text{with} \qquad \boldsymbol{Z}_k^{(n)} := \begin{bmatrix} n^{-1}\boldsymbol{M}_k \\ n^{-3/2}\langle \boldsymbol{1}, \boldsymbol{M}_k\rangle U_{k-1} \\ n^{-3/2}\boldsymbol{M}_k V_{k-1} \end{bmatrix}$$

for $t\in\mathbb{R}_+$ and $k,n\in\mathbb{N}$. Theorem 4.2 follows from Lemma A.2 and the following theorem (this will be explained after Theorem 6.1).

6.1 Theorem. *If* $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\boldsymbol{1}, \boldsymbol{1}\rangle = 0$ *then*

$$\boldsymbol{\mathcal{Z}}^{(n)} \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{Z}} \qquad \text{as } n\to\infty, \tag{6.1}$$

where the process $(\boldsymbol{\mathcal{Z}}_t)_{t\in\mathbb{R}_+}$ *with values in* $\mathbb{R}^2\times\mathbb{R}\times\mathbb{R}^2$ *is the unique strong solution of the SDE*

$$\mathrm{d}\boldsymbol{\mathcal{Z}}_t = \gamma(t, \boldsymbol{\mathcal{Z}}_t) \begin{bmatrix} \mathrm{d}\boldsymbol{\mathcal{W}}_t \\ \mathrm{d}\widetilde{\widetilde{\mathcal{W}}}_t \\ \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_t \end{bmatrix}, \qquad t\in\mathbb{R}_+, \tag{6.2}$$

with initial value $\boldsymbol{\mathcal{Z}}_0 = \boldsymbol{0}$, *where* $(\boldsymbol{\mathcal{W}}_t)_{t\in\mathbb{R}_+}$, $(\widetilde{\widetilde{\mathcal{W}}}_t)_{t\in\mathbb{R}_+}$ *and* $(\widetilde{\boldsymbol{\mathcal{W}}}_t)_{t\in\mathbb{R}_+}$ *are independent standard Wiener processes of dimension 2, 1 and 2, respectively, and the function* $\gamma:\mathbb{R}_+\times(\mathbb{R}^2\times\mathbb{R}\times\mathbb{R}^2)\to\mathbb{R}^{5\times5}$ *is defined by*

$$\gamma(t,\boldsymbol{x}) := \begin{bmatrix} \langle \boldsymbol{1}, (\boldsymbol{x}_1 + t\boldsymbol{m}_{\boldsymbol{\varepsilon}})^+\rangle^{1/2}\,\overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \langle \boldsymbol{V}_{\boldsymbol{\varepsilon}}\boldsymbol{1}, \boldsymbol{1}\rangle^{1/2}\,\langle \boldsymbol{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle\, t & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \left(\frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{u}}\rangle}{4\alpha\beta}\right)^{1/2} \langle \boldsymbol{1}, \boldsymbol{x}_1 + t\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \end{bmatrix}$$

for $t\in\mathbb{R}_+$ *and* $\boldsymbol{x} = (\boldsymbol{x}_1, x_2, \boldsymbol{x}_3)\in\mathbb{R}^2\times\mathbb{R}\times\mathbb{R}^2$.

As in the case of Theorem 4.1, the SDE (6.2) has a unique strong solution with initial value $\boldsymbol{\mathcal{Z}}_0 = \mathbf{0}$, for which we have

$$\boldsymbol{\mathcal{Z}}_t = \begin{bmatrix} \boldsymbol{\mathcal{M}}_t \\ \mathcal{N}_t \\ \boldsymbol{\mathcal{P}}_t \end{bmatrix} = \begin{bmatrix} \int_0^t \mathcal{Y}_t^{1/2} \, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \, \mathrm{d}\boldsymbol{\mathcal{W}}_s \\ \langle \boldsymbol{V}_{\boldsymbol{\varepsilon}} \mathbf{1}, \mathbf{1} \rangle^{1/2} \, \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \int_0^t s \, \mathrm{d}\widetilde{\widetilde{\mathcal{W}}}_s \\ \left(\frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle}{4\alpha\beta} \right)^{1/2} \int_0^t \mathcal{Y}_t \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \, \mathrm{d}\widetilde{\boldsymbol{\mathcal{W}}}_s \end{bmatrix}, \qquad t \in \mathbb{R}_+,$$

where now $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \mathbf{1}, \mathbf{1} \rangle = 0$ yields $\mathcal{Y}_t = \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle t$, $t \in \mathbb{R}_+$. One can again easily derive

$$\begin{bmatrix} \boldsymbol{\mathcal{X}}^{(n)} \\ \boldsymbol{\mathcal{Z}}^{(n)} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \boldsymbol{\mathcal{X}} \\ \boldsymbol{\mathcal{Z}} \end{bmatrix} \qquad \text{as } n \to \infty, \tag{6.3}$$

where

$$\boldsymbol{\mathcal{X}}_t^{(n)} := n^{-1} \boldsymbol{X}_{\lfloor nt \rfloor}, \qquad \boldsymbol{\mathcal{X}}_t := \frac{1}{2} \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_t + t \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \mathbf{1} = \frac{t}{2} \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \mathbf{1}, \qquad t \in \mathbb{R}_+, \qquad n \in \mathbb{N},$$

since $\boldsymbol{\mathcal{X}}_t = \frac{1}{2} \mathcal{Y}_t \mathbf{1} = \frac{t}{2} \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \mathbf{1}$, $t \in \mathbb{R}_+$. Next, similarly to the proof of (A.6), by Lemma C.3, convergence (6.3) and Lemma A.2 with $U_{k-1} = \langle \mathbf{1}, \boldsymbol{X}_{k-1} \rangle$ imply

$$\sum_{k=1}^n \begin{bmatrix} n^{-3} U_{k-1}^2 \\ n^{-2} V_{k-1}^2 \\ n^{-3/2} \langle \mathbf{1}, \boldsymbol{M}_k \rangle U_{k-1} \\ n^{-3/2} \langle \widetilde{\boldsymbol{u}}, \boldsymbol{M}_k \rangle V_{k-1} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \int_0^1 \langle \mathbf{1}, \boldsymbol{\mathcal{X}}_t \rangle^2 \, \mathrm{d}t \\ \left(\frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle}{4\alpha\beta} \right) \int_0^1 \langle \mathbf{1}, \boldsymbol{\mathcal{X}}_t \rangle \, \mathrm{d}t \\ \langle \boldsymbol{V}_{\boldsymbol{\varepsilon}} \mathbf{1}, \mathbf{1} \rangle^{1/2} \, \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \int_0^1 t \, \mathrm{d}\widetilde{\widetilde{\mathcal{W}}}_t \\ \left(\frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle}{4\alpha\beta} \right)^{1/2} \int_0^1 \mathcal{Y}_t \, \mathrm{d} \langle \widetilde{\boldsymbol{u}}, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \widetilde{\boldsymbol{\mathcal{W}}}_t \rangle \end{bmatrix} \qquad \text{as } n \to \infty.$$

This limiting random vector can be written in the form as given in Theorem 4.2 since $\langle \mathbf{1}, \boldsymbol{\mathcal{X}}_t \rangle = \mathcal{Y}_t = \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \, t$, and $\langle \widetilde{\boldsymbol{u}}, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} \widetilde{\boldsymbol{\mathcal{W}}}_t \rangle = \langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle^{1/2} \, \widetilde{\mathcal{W}}_t$ for all $t \in \mathbb{R}_+$ with a (one-dimensional) standard Wiener process $(\widetilde{\mathcal{W}}_t)_{t \in \mathbb{R}_+}$.

Proof of Theorem 6.1. Similar to the proof of Theorem 5.1. The conditional variance has the form

$$\mathbb{E}\big(\boldsymbol{Z}_k^{(n)} (\boldsymbol{Z}_k^{(n)})^\top \,|\, \mathcal{F}_{k-1} \big) = \begin{bmatrix} n^{-2} \boldsymbol{V}_{\boldsymbol{M}_k} & n^{-5/2} U_{k-1} \boldsymbol{V}_{\boldsymbol{M}_k} \mathbf{1} & n^{-5/2} V_{k-1} \boldsymbol{V}_{\boldsymbol{M}_k} \\ n^{-5/2} U_{k-1} \mathbf{1}^\top \boldsymbol{V}_{\boldsymbol{M}_k} & n^{-3} U_{k-1}^2 \mathbf{1}^\top \boldsymbol{V}_{\boldsymbol{M}_k} \mathbf{1} & n^{-3} U_{k-1} V_{k-1} \mathbf{1}^\top \boldsymbol{V}_{\boldsymbol{M}_k} \\ n^{-5/2} V_{k-1} \boldsymbol{V}_{\boldsymbol{M}_k} & n^{-3} U_{k-1} V_{k-1} \boldsymbol{V}_{\boldsymbol{M}_k} \mathbf{1} & n^{-3} V_{k-1}^2 \boldsymbol{V}_{\boldsymbol{M}_k} \end{bmatrix},$$

for $n \in \mathbb{N}$, $k \in \{1, \ldots, n\}$, with $\boldsymbol{V}_{\boldsymbol{M}_k} := \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \,|\, \mathcal{F}_{k-1})$, and $\gamma(s, \boldsymbol{\mathcal{Z}}_s^{(n)}) \gamma(s, \boldsymbol{\mathcal{Z}}_s^{(n)})^\top$ has the form

$$\begin{bmatrix} \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \overline{\boldsymbol{V}}_\xi & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \langle \boldsymbol{V}_{\boldsymbol{\varepsilon}} \mathbf{1}, \mathbf{1} \rangle \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^2 \, s^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle}{4\alpha\beta} \langle \mathbf{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^2 \, \overline{\boldsymbol{V}}_\xi \end{bmatrix}$$

for $s \in \mathbb{R}_+$.

In order to check condition (i) of Theorem D.1, we need to prove only that for each $T > 0$,

$$\sup_{t\in[0,T]} \left\| \frac{1}{n^{5/2}} \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1} \mathbf{1}^\top \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{6.4}$$

$$\sup_{t\in[0,T]} \left| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2 \mathbf{1}^\top \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) \mathbf{1} - \int_0^t \langle \boldsymbol{V}_{\boldsymbol{\varepsilon}} \mathbf{1}, \mathbf{1}\rangle \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^2 s^2 \, \mathrm{d}s \right| \stackrel{\mathbb{P}}{\longrightarrow} 0, \tag{6.5}$$

$$\sup_{t\in[0,T]} \left\| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1} V_{k-1} \mathbf{1}^\top \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0 \tag{6.6}$$

as $n \to \infty$, since the rest, namely, (5.5), (5.8) and (5.9) have already been proved.

Clearly, $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \mathbf{1}, \mathbf{1}\rangle = 0$ implies $\langle \boldsymbol{V}_{\boldsymbol{\xi}_1} \mathbf{1}, \mathbf{1}\rangle = 0$ and $\langle \boldsymbol{V}_{\boldsymbol{\xi}_2} \mathbf{1}, \mathbf{1}\rangle = 0$. For each $i \in \{1, 2\}$, we have $\langle \boldsymbol{V}_{\boldsymbol{\xi}_i} \mathbf{1}, \mathbf{1}\rangle = \mathbf{1}^\top \boldsymbol{V}_{\boldsymbol{\xi}_i} \mathbf{1} = (\boldsymbol{V}_{\boldsymbol{\xi}_i}^{1/2} \mathbf{1})^\top (\boldsymbol{V}_{\boldsymbol{\xi}_i}^{1/2} \mathbf{1}) = \|\boldsymbol{V}_{\boldsymbol{\xi}_i}^{1/2} \mathbf{1}\|^2$, hence we obtain $\boldsymbol{V}_{\boldsymbol{\xi}_i}^{1/2} \mathbf{1} = \mathbf{0}$, thus $\boldsymbol{V}_{\boldsymbol{\xi}_i} \mathbf{1} = \boldsymbol{V}_{\boldsymbol{\xi}_i}^{1/2} (\boldsymbol{V}_{\boldsymbol{\xi}_i}^{1/2} \mathbf{1}) = \mathbf{0}$, and hence $\mathbf{1}^\top \boldsymbol{V}_{\boldsymbol{\xi}_i} = \mathbf{0}$.

First we show (6.4). By (5.14),

$$\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1} \mathbf{1}^\top \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) = \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1} \mathbf{1}^\top \boldsymbol{V}_{\boldsymbol{\varepsilon}},$$

hence using (B.5) with $\ell = 2$, $i = 1$, $j = 0$, we conclude (6.4).

Now we turn to check (6.5). By (5.19),

$$\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2 \mathbf{1}^\top \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) \mathbf{1} = \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2 \mathbf{1}^\top \boldsymbol{V}_{\boldsymbol{\varepsilon}} \mathbf{1} = \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2 \langle \boldsymbol{V}_{\boldsymbol{\varepsilon}} \mathbf{1}, \mathbf{1}\rangle,$$

hence, in order to show (6.5), it suffices to prove

$$\sup_{t\in[0,T]} \left| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2 - \frac{t^3}{3} \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^2 \right| \stackrel{\mathbb{P}}{\longrightarrow} 0. \tag{6.7}$$

We have

$$\left| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}^2 - \frac{t^3}{3} \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^2 \right| \leqslant \frac{1}{n^3} \sum_{k=1}^{\lfloor nt\rfloor} \left| U_{k-1}^2 - (k-1)^2 \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^2 \right| + \left| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt\rfloor} (k-1)^2 - \frac{t^3}{3} \right| \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^2,$$

where

$$\sup_{t\in[0,T]} \left| \frac{1}{n^3} \sum_{k=1}^{\lfloor nt\rfloor} (k-1)^2 - \frac{t^3}{3} \right| = \sup_{t\in[0,T]} \frac{|\lfloor nt\rfloor(\lfloor nt\rfloor - 1)(2\lfloor nt\rfloor - 1) - 2n^3 t^3|}{6n^3} \to 0$$

as $n \to \infty$, hence, in order to show (6.5), it suffices to prove

$$\frac{1}{n^3} \sup_{t \in [0,T]} \sum_{k=1}^{\lfloor nt \rfloor} \big| U_k^2 - k^2 \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^2 \big| = \frac{1}{n^3} \sum_{k=1}^{\lfloor nT \rfloor} \big| U_k^2 - k^2 \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^2 \big| \xrightarrow{\mathbb{P}} 0. \tag{6.8}$$

For all $k \in \mathbb{N}$, by Remark 3.3, $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \mathbf{1}, \mathbf{1} \rangle = 0$ implies

$$\begin{aligned} U_k = X_{k,1} + X_{k,2} &= \sum_{j=1}^{X_{k-1,1}} (\xi_{k,j,1,1} + \xi_{k,j,1,2}) + \sum_{j=1}^{X_{k-1,2}} (\xi_{k,j,2,1} + \xi_{k,j,2,2}) + (\varepsilon_{k,1} + \varepsilon_{k,2}) \\ &\overset{\text{a.s.}}{=} X_{k-1,1} + X_{k-1,2} + \varepsilon_{k,1} + \varepsilon_{k,2} = U_{k-1} + \langle \mathbf{1}, \boldsymbol{\varepsilon}_k \rangle, \end{aligned}$$

hence $U_k = \sum_{i=1}^k \langle \mathbf{1}, \boldsymbol{\varepsilon}_i \rangle$. Applying Kolmogorov's maximal inequality, we obtain

$$\mathbb{P}\left(n^{-1} \max_{k \in \{1,\ldots,\lfloor nT \rfloor\}} |U_k - k \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle| \geqslant \varepsilon \right) \leqslant n^{-2} \varepsilon^{-2} \operatorname{Var}(U_{\lfloor nT \rfloor}) = \frac{\lfloor nT \rfloor}{n^2 \varepsilon^2} \operatorname{Var}(\langle \mathbf{1}, \boldsymbol{\varepsilon}_1 \rangle^2) \to 0$$

as $n \to \infty$ for all $\varepsilon > 0$, thus

$$n^{-1} \max_{k \in \{1,\ldots,\lfloor nT \rfloor\}} |U_k - k \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty.$$

We have

$$|U_k^2 - k^2 \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^2| \leqslant |U_k - k \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle|^2 + 2k \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle |U_k - k \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle|,$$

hence

$$\begin{aligned} &n^{-2} \max_{k \in \{1,\ldots,\lfloor nT \rfloor\}} |U_k^2 - k^2 \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^2| \\ &\leqslant \left(n^{-1} \max_{k \in \{1,\ldots,\lfloor nT \rfloor\}} |U_k - k \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle| \right)^2 + \frac{2\lfloor nT \rfloor}{n^2} \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \max_{k \in \{1,\ldots,\lfloor nT \rfloor\}} |U_k - k \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle| \xrightarrow{\mathbb{P}} 0 \end{aligned}$$

as $n \to \infty$. Consequently,

$$\frac{1}{n^3} \sum_{k=1}^{\lfloor nT \rfloor} \big| U_{k-1}^2 - (k-1)^2 \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^2 \big| \leqslant \frac{\lfloor nT \rfloor}{n^3} \max_{k \in \{1,\ldots,\lfloor nT \rfloor\}} \big| U_{k-1}^2 - (k-1)^2 \langle \mathbf{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^2 \big| \xrightarrow{\mathbb{P}} 0$$

as $n \to \infty$, thus we conclude (6.8), and hence (6.5).

Finally, we check (6.6). By (5.40),

$$\sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1} \mathbf{1}^\top \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) = \sum_{k=1}^{\lfloor nt \rfloor} U_{k-1} V_{k-1} \mathbf{1}^\top \boldsymbol{V}_{\boldsymbol{\varepsilon}},$$

hence using (B.5) with $\ell = 2$, $i = 1$, $j = 1$, we conclude (6.6).

Condition (ii) of Theorem D.1 can be checked as in case of Theorem 5.1. $\square$

7 Proof of Theorem 4.3

This proof is also similar to the proof of Theorem 4.1. Consider the sequence of stochastic processes

$$\boldsymbol{\mathcal{Z}}_t^{(n)} := \begin{bmatrix} \boldsymbol{\mathcal{M}}_t^{(n)} \\ \boldsymbol{\mathcal{N}}_t^{(n)} \\ \mathcal{P}_t^{(n)} \end{bmatrix} := \sum_{k=1}^{\lfloor nt \rfloor} \boldsymbol{Z}_k^{(n)} \qquad \text{with} \qquad \boldsymbol{Z}_k^{(n)} := \begin{bmatrix} n^{-1}\boldsymbol{M}_k \\ n^{-2}\boldsymbol{M}_k U_{k-1} \\ n^{-1/2}\langle \widetilde{\boldsymbol{u}}, \boldsymbol{M}_k\rangle V_{k-1} \end{bmatrix}$$

for $t \in \mathbb{R}_+$ and $k, n \in \mathbb{N}$. Theorem 4.3 follows from Lemma A.3 and the following theorem (this will be explained after Theorem 7.1).

7.1 Theorem. *If* $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle = 0$ *then*

$$\boldsymbol{\mathcal{Z}}^{(n)} \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{Z}} \qquad \text{as } n \to \infty, \tag{7.1}$$

where the process $(\boldsymbol{\mathcal{Z}}_t)_{t\in\mathbb{R}_+}$ *with values in* $\mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}$ *is the unique strong solution of the SDE*

$$\mathrm{d}\boldsymbol{\mathcal{Z}}_t = \gamma(t, \boldsymbol{\mathcal{Z}}_t) \begin{bmatrix} \mathrm{d}\boldsymbol{\mathcal{W}}_t \\ \mathrm{d}\widetilde{\mathcal{W}}_t \end{bmatrix}, \qquad t \in \mathbb{R}_+, \tag{7.2}$$

with initial value $\boldsymbol{\mathcal{Z}}_0 = \boldsymbol{0}$, *where* $(\boldsymbol{\mathcal{W}}_t)_{t\in\mathbb{R}_+}$ *and* $(\widetilde{\mathcal{W}}_t)_{t\in\mathbb{R}_+}$ *are independent standard Wiener processes of dimension 2 and 1, respectively, and* $\gamma : \mathbb{R}_+ \times (\mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}) \to \mathbb{R}^{5\times 3}$ *is defined by*

$$\gamma(t, \boldsymbol{x}) := \begin{bmatrix} \langle \boldsymbol{1}, (\boldsymbol{x}_1 + t\boldsymbol{m}_{\boldsymbol{\varepsilon}})^+\rangle^{1/2}\, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} & \boldsymbol{0} \\ \langle \boldsymbol{1}, (\boldsymbol{x}_1 + t\boldsymbol{m}_{\boldsymbol{\varepsilon}})^+\rangle^{3/2}\, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2} & \boldsymbol{0} \\ \boldsymbol{0} & [\langle \boldsymbol{V}_{\boldsymbol{\varepsilon}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle\, \mathbb{E}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_1\rangle^2)]^{1/2} \end{bmatrix}$$

for $t \in \mathbb{R}_+$ *and* $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, x_3) \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}$.

As in the case of Theorem 4.1, the SDE (7.2) has a unique strong solution with initial value $\boldsymbol{\mathcal{Z}}_0 = \boldsymbol{0}$, for which we have

$$\boldsymbol{\mathcal{Z}}_t = \begin{bmatrix} \boldsymbol{\mathcal{M}}_t \\ \boldsymbol{\mathcal{N}}_t \\ \mathcal{P}_t \end{bmatrix} = \begin{bmatrix} \int_0^t \mathcal{Y}_s^{1/2}\, \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}^{1/2}\, \mathrm{d}\boldsymbol{\mathcal{W}}_s \\ \int_0^t \mathcal{Y}_s\, \mathrm{d}\boldsymbol{\mathcal{M}}_s \\ [\langle \boldsymbol{V}_{\boldsymbol{\varepsilon}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle\, \mathbb{E}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_1\rangle^2)]^{1/2}\, \widetilde{\mathcal{W}}_t \end{bmatrix}, \qquad t \in \mathbb{R}_+.$$

One can again easily derive

$$\begin{bmatrix} \boldsymbol{\mathcal{X}}^{(n)} \\ \boldsymbol{\mathcal{Z}}^{(n)} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \boldsymbol{\mathcal{X}} \\ \boldsymbol{\mathcal{Z}} \end{bmatrix} \qquad \text{as } n \to \infty, \tag{7.3}$$

where
$$\boldsymbol{\mathcal{X}}_t^{(n)} := n^{-1}\boldsymbol{X}_{\lfloor nt\rfloor}, \qquad \boldsymbol{\mathcal{X}}_t := \frac{1}{2}\langle \boldsymbol{1}, \boldsymbol{\mathcal{M}}_t + t\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle \boldsymbol{1}, \qquad t \in \mathbb{R}_+, \qquad n \in \mathbb{N}.$$
Next, similarly to the proof of (A.6), by Lemma C.3, convergence (7.3) and Lemma A.3 imply
$$\sum_{k=1}^n \begin{bmatrix} n^{-3}U_{k-1}^2 \\ n^{-1}V_{k-1}^2 \\ n^{-2}\langle \boldsymbol{1}, \boldsymbol{M}_k\rangle U_{k-1} \\ n^{-1/2}\langle \widetilde{\boldsymbol{u}}, \boldsymbol{M}_k\rangle V_{k-1} \end{bmatrix} \xrightarrow{\mathcal{D}} \begin{bmatrix} \int_0^1 \langle \boldsymbol{1}, \boldsymbol{\mathcal{X}}_t\rangle^2 \,\mathrm{d}t \\ \mathbb{E}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_1\rangle^2) \\ \int_0^1 \mathcal{Y}_t \,\mathrm{d}\langle \boldsymbol{1}, \boldsymbol{\mathcal{M}}_t\rangle \\ [\langle \boldsymbol{V}_{\boldsymbol{\varepsilon}}\widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle \,\mathbb{E}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_1\rangle^2)]^{1/2}\, \widetilde{\mathcal{W}}_1 \end{bmatrix} \qquad \text{as } n \to \infty.$$
Note that this convergence holds even in case $\mathbb{E}[\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_1\rangle^2] = 0$. The limiting random vector can be written in the form as given in Theorem 4.3, since $\langle \boldsymbol{1}, \boldsymbol{\mathcal{X}}_t\rangle = \mathcal{Y}_t$ and $\langle \boldsymbol{1}, \boldsymbol{\mathcal{M}}_t\rangle = \mathcal{Y}_t - \langle \boldsymbol{1}, \boldsymbol{m}_{\boldsymbol{\varepsilon}}\, t\rangle$ for all $t \in \mathbb{R}_+$.

Proof of Theorem 7.1. Similar to the proof of Theorem 5.1. The conditional variance $\mathbb{E}\big(\boldsymbol{Z}_k^{(n)}(\boldsymbol{Z}_k^{(n)})^\top \,|\, \mathcal{F}_{k-1}\big)$ has the form
$$\begin{bmatrix} n^{-2}\boldsymbol{V}_{\boldsymbol{M}_k} & n^{-3}U_{k-1}\boldsymbol{V}_{\boldsymbol{M}_k} & n^{-3/2}V_{k-1}\boldsymbol{V}_{\boldsymbol{M}_k}\widetilde{\boldsymbol{u}} \\ n^{-3}U_{k-1}\boldsymbol{V}_{\boldsymbol{M}_k} & n^{-4}U_{k-1}^2\boldsymbol{V}_{\boldsymbol{M}_k} & n^{-5/2}U_{k-1}V_{k-1}\boldsymbol{V}_{\boldsymbol{M}_k}\widetilde{\boldsymbol{u}} \\ n^{-3/2}V_{k-1}\widetilde{\boldsymbol{u}}^\top\boldsymbol{V}_{\boldsymbol{M}_k} & n^{-5/2}U_{k-1}V_{k-1}\widetilde{\boldsymbol{u}}^\top\boldsymbol{V}_{\boldsymbol{M}_k} & n^{-1}V_{k-1}^2\widetilde{\boldsymbol{u}}^\top\boldsymbol{V}_{\boldsymbol{M}_k}\widetilde{\boldsymbol{u}} \end{bmatrix}$$
for $n \in \mathbb{N}$, $k \in \{1, \ldots, n\}$, with $\boldsymbol{V}_{\boldsymbol{M}_k} := \mathbb{E}(\boldsymbol{M}_k\boldsymbol{M}_k^\top \,|\, \mathcal{F}_{k-1})$, and $\gamma(s, \boldsymbol{\mathcal{Z}}_s^{(n)})\gamma(s, \boldsymbol{\mathcal{Z}}_s^{(n)})^\top$ has the form
$$\begin{bmatrix} \langle \boldsymbol{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle \overline{\boldsymbol{V}}_{\xi} & \langle \boldsymbol{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^2 \overline{\boldsymbol{V}}_{\xi} & \boldsymbol{0} \\ \langle \boldsymbol{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^2 \overline{\boldsymbol{V}}_{\xi} & \langle \boldsymbol{1}, \boldsymbol{\mathcal{M}}_s^{(n)} + s\boldsymbol{m}_{\boldsymbol{\varepsilon}}\rangle^3 \overline{\boldsymbol{V}}_{\xi} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \langle \boldsymbol{V}_{\boldsymbol{\varepsilon}}\widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle \,\mathbb{E}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_1\rangle^2) \end{bmatrix}$$
for $s \in \mathbb{R}_+$.

In order to check condition (i) of Theorem D.1, we need to prove only that for each $T > 0$,

$$\sup_{t\in[0,T]} \left| \frac{1}{n} \sum_{k=1}^{\lfloor nt\rfloor} V_{k-1}^2 \widetilde{\boldsymbol{u}}^\top \mathbb{E}(\boldsymbol{M}_k\boldsymbol{M}_k^\top \,|\, \mathcal{F}_{k-1})\widetilde{\boldsymbol{u}} - t\langle \boldsymbol{V}_{\boldsymbol{\varepsilon}}\widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle \,\mathbb{E}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_1\rangle^2) \right| \xrightarrow{\mathbb{P}} 0, \tag{7.4}$$

$$\sup_{t\in[0,T]} \left\| \frac{1}{n^{3/2}} \sum_{k=1}^{\lfloor nt\rfloor} V_{k-1} \widetilde{\boldsymbol{u}}^\top \mathbb{E}(\boldsymbol{M}_k\boldsymbol{M}_k^\top \,|\, \mathcal{F}_{k-1}) \right\| \xrightarrow{\mathbb{P}} 0, \tag{7.5}$$

$$\sup_{t\in[0,T]} \left\| \frac{1}{n^{5/2}} \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}V_{k-1} \widetilde{\boldsymbol{u}}^\top \mathbb{E}(\boldsymbol{M}_k\boldsymbol{M}_k^\top \,|\, \mathcal{F}_{k-1}) \right\| \xrightarrow{\mathbb{P}} 0 \tag{7.6}$$

as $n \to \infty$, since the rest, namely, (5.5), (5.6) and (5.7), have already been proved.

Clearly, $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}}\widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle = 0$ implies $\langle \boldsymbol{V}_{\boldsymbol{\xi}_1}\widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle = 0$ and $\langle \boldsymbol{V}_{\boldsymbol{\xi}_2}\widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle = 0$. For each $i \in \{1, 2\}$, we have $\langle \boldsymbol{V}_{\boldsymbol{\xi}_i}\widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle = \widetilde{\boldsymbol{u}}^\top\boldsymbol{V}_{\boldsymbol{\xi}_i}\widetilde{\boldsymbol{u}} = (\boldsymbol{V}_{\boldsymbol{\xi}_i}^{1/2}\widetilde{\boldsymbol{u}})^\top(\boldsymbol{V}_{\boldsymbol{\xi}_i}^{1/2}\widetilde{\boldsymbol{u}}) = \|\boldsymbol{V}_{\boldsymbol{\xi}_i}^{1/2}\widetilde{\boldsymbol{u}}\|^2$, hence we obtain $\boldsymbol{V}_{\boldsymbol{\xi}_i}^{1/2}\widetilde{\boldsymbol{u}} = \boldsymbol{0}$, thus $\boldsymbol{V}_{\boldsymbol{\xi}_i}\widetilde{\boldsymbol{u}} = \boldsymbol{V}_{\boldsymbol{\xi}_i}^{1/2}(\boldsymbol{V}_{\boldsymbol{\xi}_i}^{1/2}\widetilde{\boldsymbol{u}}) = \boldsymbol{0}$, and hence $\widetilde{\boldsymbol{u}}^\top\boldsymbol{V}_{\boldsymbol{\xi}_i} = \boldsymbol{0}$.

First we show (7.4). By (5.27),

$$\sum_{k=1}^{\lfloor nt\rfloor} V_{k-1}^2 \widetilde{\boldsymbol{u}}^\top \operatorname{\mathbb{E}}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) \widetilde{\boldsymbol{u}} = \sum_{k=1}^{\lfloor nt\rfloor} V_{k-1}^2 \widetilde{\boldsymbol{u}}^\top \boldsymbol{V}_{\boldsymbol{\varepsilon}} \widetilde{\boldsymbol{u}},$$

hence, in order to show (7.4), it suffices to prove

$$\sup_{t\in[0,T]} \left| \frac{1}{n} \sum_{k=1}^{\lfloor nt\rfloor} V_{k-1}^2 - t \operatorname{\mathbb{E}}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_1\rangle^2) \right| \stackrel{\mathbb{P}}{\longrightarrow} 0.$$

For all $k \in \mathbb{N}$, by Remark 3.3, $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle = 0$ implies

$$\begin{aligned}
V_k = X_{k,1} - X_{k,2} &= \sum_{j=1}^{X_{k-1,1}} (\xi_{k,j,1,1} - \xi_{k,j,1,2}) + \sum_{j=1}^{X_{k-1,2}} (\xi_{k,j,2,1} - \xi_{k,j,2,2}) + (\varepsilon_{k,1} - \varepsilon_{k,2}) \\
&\stackrel{\text{a.s.}}{=} \varepsilon_{k,1} - \varepsilon_{k,2} = \langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_k\rangle.
\end{aligned}$$

We have

$$\left| \frac{1}{n} \sum_{k=1}^{\lfloor nt\rfloor} V_{k-1}^2 - t \operatorname{\mathbb{E}}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_1\rangle^2) \right| \leqslant \frac{1}{n} \left| \sum_{k=1}^{\lfloor nt\rfloor} \big[\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_{k-1}\rangle^2 - \operatorname{\mathbb{E}}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_{k-1}\rangle^2) \big] \right| + \frac{|nt - \lfloor nt\rfloor|}{n} \operatorname{\mathbb{E}}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_k\rangle^2),$$

where $|nt - \lfloor nt\rfloor| \leqslant 1$, hence, in order to show (7.4), it suffices to prove

$$\frac{1}{n} \sup_{t\in[0,T]} \left| \sum_{k=1}^{\lfloor nt\rfloor} \big[\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_k\rangle^2 - \operatorname{\mathbb{E}}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_k\rangle^2) \big] \right| = \frac{1}{n} \max_{N\in\{1,\dots,\lfloor nT\rfloor\}} \left| \sum_{k=1}^{N} \big[\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_k\rangle^2 - \operatorname{\mathbb{E}}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_k\rangle^2) \big] \right| \stackrel{\mathbb{P}}{\longrightarrow} 0. \tag{7.7}$$

Applying Kolmogorov's maximal inequality, we obtain

$$\begin{aligned}
&\mathbb{P}\left(n^{-1} \max_{N\in\{1,\dots,\lfloor nT\rfloor\}} \left| \sum_{k=1}^{N} \big[\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_k\rangle^2 - \operatorname{\mathbb{E}}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_k\rangle^2) \big] \right| \geqslant \varepsilon \right) \\
&\qquad\qquad \leqslant \frac{1}{n^2\varepsilon^2} \operatorname{Var}\left(\sum_{k=1}^{\lfloor nT\rfloor} \langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_k\rangle^2 \right) = \frac{\lfloor nT\rfloor}{n^2\varepsilon^2} \operatorname{Var}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_k\rangle^2) \to 0 \qquad \text{as } n \to \infty
\end{aligned}$$

for all $\varepsilon > 0$, thus we conclude (7.7), and hence (7.4).

Now we turn to check (7.5). By (5.35),

$$\sum_{k=1}^{\lfloor nt\rfloor} V_{k-1} \widetilde{\boldsymbol{u}}^\top \operatorname{\mathbb{E}}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) = \sum_{k=1}^{\lfloor nt\rfloor} V_{k-1} \widetilde{\boldsymbol{u}}^\top \boldsymbol{V}_{\boldsymbol{\varepsilon}}.$$

Again by the strong law of large numbers, $n^{-1} \sum_{k=1}^{\lfloor nT\rfloor} |V_{k-1}| \stackrel{\text{a.s.}}{\longrightarrow} t \operatorname{\mathbb{E}}(|\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_1\rangle|)$ as $n \to \infty$ for all $T > 0$, hence we conclude (7.5).

Finally, we check (7.6). By (5.40),

$$\sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}V_{k-1}\widetilde{\boldsymbol{u}}^\top \operatorname{\mathbb{E}}(\boldsymbol{M}_k\boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) = \sum_{k=1}^{\lfloor nt\rfloor} U_{k-1}V_{k-1}\widetilde{\boldsymbol{u}}^\top \boldsymbol{V}_{\boldsymbol{\varepsilon}}.$$

Applying again $V_k = \langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_k\rangle$, $k \in \mathbb{N}$, and Corollary B.7, we obtain $\operatorname{\mathbb{E}}(|U_{k-1}V_{k-1}|) \leqslant \sqrt{\operatorname{\mathbb{E}}(U_{k-1}^2)\operatorname{\mathbb{E}}(V_{k-1}^2)} = \mathrm{O}(k)$, which clearly implies (7.6).

Condition (ii) of Theorem D.1 can be checked again as in case of Theorem 5.1. □

Appendices

A CLS estimators

In order to find CLS estimators of the criticality parameter $\varrho = \alpha + \beta$, we introduce a further parameter $\delta := \alpha - \beta$. Then $\alpha = (\varrho + \delta)/2$ and $\beta = (\varrho - \delta)/2$, thus the recursion (4.2) can be written in the form

$$\boldsymbol{X}_k = \frac{1}{2}\begin{bmatrix} \varrho+\delta & \varrho-\delta \\ \varrho-\delta & \varrho+\delta \end{bmatrix} \boldsymbol{X}_{k-1} + \boldsymbol{M}_k + \boldsymbol{m}_{\boldsymbol{\varepsilon}}, \qquad k \in \mathbb{N}.$$

For each $n \in \mathbb{N}$, a CLS estimator $(\widehat{\varrho}_n(\boldsymbol{X}_1,\ldots,\boldsymbol{X}_n), \widehat{\delta}_n(\boldsymbol{X}_1,\ldots,\boldsymbol{X}_n))$ of (ϱ,δ) based on a sample $\boldsymbol{X}_1,\ldots,\boldsymbol{X}_n$ can be obtained by minimizing the sum of squares

$$\sum_{k=1}^n \big\|\boldsymbol{X}_k - \operatorname{\mathbb{E}}(\boldsymbol{X}_k \mid \mathcal{F}_{k-1})\big\|^2 = \sum_{k=1}^n \left\| \boldsymbol{X}_k - \frac{1}{2}\begin{bmatrix} \varrho+\delta & \varrho-\delta \\ \varrho-\delta & \varrho+\delta \end{bmatrix} \boldsymbol{X}_{k-1} - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \right\|^2$$

with respect to (ϱ,δ) over $\mathbb{R}^2$. In what follows, we use the notation $\boldsymbol{x}_0 := \boldsymbol{0}$. For all $n \in \mathbb{N}$, we define the function $Q_n : (\mathbb{R}^2)^n \times \mathbb{R}^2 \to \mathbb{R}$ by

$$Q_n(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n;\varrho',\delta') := \sum_{k=1}^n \left\| \boldsymbol{x}_k - \frac{1}{2}\begin{bmatrix} \varrho'+\delta' & \varrho'-\delta' \\ \varrho'-\delta' & \varrho'+\delta' \end{bmatrix} \boldsymbol{x}_{k-1} - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \right\|^2$$

for all $(\varrho',\delta') \in \mathbb{R}^2$ and $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n \in \mathbb{R}^2$. By definition, for all $n \in \mathbb{N}$, a CLS estimator of (ϱ,δ) is a measurable function $(\widehat{\varrho}_n,\widehat{\delta}_n) : (\mathbb{R}^2)^n \to \mathbb{R}^2$ such that

$$Q_n(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n;\widehat{\varrho}_n(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n),\widehat{\delta}_n(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)) = \inf_{(\varrho',\delta')\in\mathbb{R}^2} Q_n(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n;\varrho',\delta')$$

for all $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n \in \mathbb{R}^2$. Next we give the solutions of this extremum problem.

A.1 Lemma. *For each* $n \in \mathbb{N}$, *any CLS estimator of* (ϱ, δ) *is a measurable function* $(\widehat{\varrho}_n, \widehat{\delta}_n) : (\mathbb{R}^2)^n \to \mathbb{R}^2$ *for which*

$$\widehat{\varrho}_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) := \frac{\sum_{k=1}^n \langle \boldsymbol{1}, \boldsymbol{x}_k - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \langle \boldsymbol{1}, \boldsymbol{x}_{k-1} \rangle}{\sum_{k=1}^n \langle \boldsymbol{1}, \boldsymbol{x}_{k-1} \rangle^2}, \tag{A.1}$$

$$\widehat{\delta}_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) := \frac{\sum_{k=1}^n \langle \widetilde{\boldsymbol{u}}, \boldsymbol{x}_k - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \langle \widetilde{\boldsymbol{u}}, \boldsymbol{x}_{k-1} \rangle}{\sum_{k=1}^n \langle \widetilde{\boldsymbol{u}}, \boldsymbol{x}_{k-1} \rangle^2} \tag{A.2}$$

on the set $H_n \cap \widetilde{H}_n$ *given in* (3.2) *and* (3.4).

Observe that (A.1) and (A.2) give natural CLS estimators of ϱ and δ on the set H_n and $\widetilde{H}_n$, respectively.

Proof of Lemma A.1. The quadratic function Q_n can be written in the form

$$\begin{aligned}
Q_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n; \varrho', \delta') &= \sum_{k=1}^n \left(\langle \boldsymbol{e}_1, \boldsymbol{x}_k - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle - \frac{1}{2}\varrho' \langle \boldsymbol{1}, \boldsymbol{x}_{k-1} \rangle - \frac{1}{2}\delta' \langle \widetilde{\boldsymbol{u}}, \boldsymbol{x}_{k-1} \rangle \right)^2 \\
&\quad + \sum_{k=1}^n \left(\langle \boldsymbol{e}_2, \boldsymbol{x}_k - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle - \frac{1}{2}\varrho' \langle \boldsymbol{1}, \boldsymbol{x}_{k-1} \rangle + \frac{1}{2}\delta' \langle \widetilde{\boldsymbol{u}}, \boldsymbol{x}_{k-1} \rangle \right)^2 \\
&= \frac{1}{2}(\varrho')^2 \sum_{k=1}^n \langle \boldsymbol{1}, \boldsymbol{x}_{k-1} \rangle^2 - \varrho' \sum_{k=1}^n \langle \boldsymbol{1}, \boldsymbol{x}_k - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \langle \boldsymbol{1}, \boldsymbol{x}_{k-1} \rangle + \sum_{k=1}^n \langle \boldsymbol{e}_1, \boldsymbol{x}_k - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^2 \\
&\quad + \frac{1}{2}(\delta')^2 \sum_{k=1}^n \langle \widetilde{\boldsymbol{u}}, \boldsymbol{x}_{k-1} \rangle^2 - \delta' \sum_{k=1}^n \langle \widetilde{\boldsymbol{u}}, \boldsymbol{x}_k - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \langle \widetilde{\boldsymbol{u}}, \boldsymbol{x}_{k-1} \rangle + \sum_{k=1}^n \langle \boldsymbol{e}_2, \boldsymbol{x}_k - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^2,
\end{aligned}$$

where $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ denote the standard basis in $\mathbb{R}^2$, hence we obtain (A.1) and (A.2). $\square$

One can easily check that any CLS estimator $(\widehat{\alpha}_n, \widehat{\beta}_n) : (\mathbb{R}^2)^n \to \mathbb{R}^2$ of (α, β) is of the form

$$\begin{bmatrix} \widehat{\alpha}_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \\ \widehat{\beta}_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \widehat{\varrho}_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \\ \widehat{\delta}_n(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \end{bmatrix}, \qquad \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathbb{R}^2. \tag{A.3}$$

Namely, if $\psi : \mathbb{R}^2 \to \mathbb{R}^2$ is the bijective measurable function such that

$$\mathbb{R}^2 \ni (\alpha', \beta') \mapsto \psi(\alpha', \beta') := \begin{bmatrix} \alpha' + \beta' \\ \alpha' - \beta' \end{bmatrix} = \begin{bmatrix} \varrho' \\ \delta' \end{bmatrix},$$

then there is a bijection between the set of CLS estimators of the parameters (α, β) and the set of CLS estimators of the parameters $\psi(\alpha, \beta)$. Indeed, for all $n \in \mathbb{N}$, $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathbb{R}^2$ and $(\alpha', \beta') \in \mathbb{R}^2$,

$$\sum_{k=1}^n \left\| \boldsymbol{x}_k - \begin{bmatrix} \alpha' & \beta' \\ \beta' & \alpha' \end{bmatrix} \boldsymbol{x}_{k-1} - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \right\|^2 = \sum_{k=1}^n \left\| \boldsymbol{x}_k - \begin{bmatrix} \psi^{-1}(\varrho', \gamma')_1 & \psi^{-1}(\varrho', \gamma')_2 \\ \psi^{-1}(\varrho', \gamma')_2 & \psi^{-1}(\varrho', \gamma')_1 \end{bmatrix} \boldsymbol{x}_{k-1} - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \right\|^2,$$

hence $(\widehat{\alpha}_n, \widehat{\beta}_n) : \mathbb{R}^n \to \mathbb{R}^2$ is a CLS estimator of (α, β) if and only if $\psi(\widehat{\alpha}_n, \widehat{\beta}_n)$ is a CLS estimator of $\psi(\alpha, \beta)$, and we obtain (A.3). Hence this CLS estimator has the form

$$\begin{bmatrix} \widehat{\alpha}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \\ \widehat{\beta}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \end{bmatrix} = A_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)^{-1} b_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \tag{A.4}$$

on the set $H_n \cap \widetilde{H}_n$, where

$$A_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) := \sum_{k=1}^n \begin{bmatrix} x_{k-1,1} & x_{k-1,2} \\ x_{k-1,2} & x_{k-1,1} \end{bmatrix}^2,$$

$$b_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) := \sum_{k=1}^n \begin{bmatrix} x_{k-1,1} & x_{k-1,2} \\ x_{k-1,2} & x_{k-1,1} \end{bmatrix} (\boldsymbol{x}_k - \boldsymbol{m}_{\boldsymbol{\varepsilon}}).$$

Indeed, by Lemma A.1,

$$\begin{bmatrix} \widehat{\alpha}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \\ \widehat{\beta}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \widehat{\varrho}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \\ \widehat{\delta}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} F_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)^{-1} g_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n),$$

where

$$F_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) := \sum_{k=1}^n \begin{bmatrix} \langle \boldsymbol{1}, \boldsymbol{x}_{k-1} \rangle^2 & 0 \\ 0 & \langle \widetilde{\boldsymbol{u}}, \boldsymbol{x}_{k-1} \rangle^2 \end{bmatrix},$$

$$g_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) := \sum_{k=1}^n \begin{bmatrix} \langle \boldsymbol{1}, \boldsymbol{x}_{k-1} \rangle & 0 \\ 0 & \langle \widetilde{\boldsymbol{u}}, \boldsymbol{x}_{k-1} \rangle \end{bmatrix} \begin{bmatrix} \langle \boldsymbol{1}, \boldsymbol{x}_k - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \\ \langle \widetilde{\boldsymbol{u}}, \boldsymbol{x}_k - \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \end{bmatrix},$$

hence

$$\begin{aligned} \begin{bmatrix} \widehat{\alpha}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \\ \widehat{\beta}_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \end{bmatrix} &= \frac{1}{2} \left(\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} F_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} \right)^{-1} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} g_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \\ &= A_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)^{-1} b_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n), \end{aligned}$$

which also shows the existence of $A_n(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)^{-1}$ on the set $H_n \cap \widetilde{H}_n$.

In order to analyse existence and uniqueness of these estimators in case of a critical doubly symmetric 2-type Galton–Watson process, i.e., when $\varrho = 1$, we need the following approximations.

A.2 Lemma. *We have*

$$n^{-2} \left(\sum_{k=1}^n V_k^2 - \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle}{4\alpha\beta} \sum_{k=1}^n U_{k-1} \right) \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty.$$

Proof. In order to prove the satement, we derive a decomposition of $\sum_{k=1}^n V_k^2$ as a sum of a martingale and some negligible terms. Using recursion (4.4), Lemma B.1 and (4.5), we obtain

$$\begin{aligned}
\mathbb{E}(V_k^2 \,|\, \mathcal{F}_{k-1}) &= \mathbb{E}\left[((\alpha-\beta)V_{k-1} + \langle \widetilde{\boldsymbol{u}}, \boldsymbol{M}_k + \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle)^2 \,\big|\, \mathcal{F}_{k-1} \right] \\
&= (\alpha-\beta)^2 V_{k-1}^2 + 2(\alpha-\beta)\langle \widetilde{\boldsymbol{u}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle V_{k-1} + \langle \widetilde{\boldsymbol{u}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle^2 + \widetilde{\boldsymbol{u}}^\top \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \,|\, \mathcal{F}_{k-1}) \widetilde{\boldsymbol{u}} \\
&= (\alpha-\beta)^2 V_{k-1}^2 + \widetilde{\boldsymbol{u}}^\top (X_{k-1,1} \boldsymbol{V}_{\boldsymbol{\xi}_1} + X_{k-1,2} \boldsymbol{V}_{\boldsymbol{\xi}_2}) \widetilde{\boldsymbol{u}} + \text{constant} + \text{constant} \times V_{k-1} \\
&= (\alpha-\beta)^2 V_{k-1}^2 + \frac{1}{2} \widetilde{\boldsymbol{u}}^\top (\boldsymbol{V}_{\boldsymbol{\xi}_1} + \boldsymbol{V}_{\boldsymbol{\xi}_2}) \widetilde{\boldsymbol{u}} \, U_{k-1} + \text{constant} + \text{constant} \times V_{k-1}.
\end{aligned}$$

Thus

$$\begin{aligned}
\sum_{k=1}^n V_k^2 &= \sum_{k=1}^n \left[V_k^2 - \mathbb{E}(V_k^2 \,|\, \mathcal{F}_{k-1}) \right] + \sum_{k=1}^n \mathbb{E}(V_k^2 \,|\, \mathcal{F}_{k-1}) \\
&= \sum_{k=1}^n \left[V_k^2 - \mathbb{E}(V_k^2 \,|\, \mathcal{F}_{k-1}) \right] + (\alpha-\beta)^2 \sum_{k=1}^n V_{k-1}^2 + \widetilde{\boldsymbol{u}}^\top \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}} \sum_{k=1}^n U_{k-1} \\
&\quad + \mathrm{O}(n) + \text{constant} \times \sum_{k=1}^n V_{k-1}.
\end{aligned}$$

Consequently,

$$\tag{A.5} \begin{aligned}
\sum_{k=1}^n V_k^2 &= \frac{1}{1-(\alpha-\beta)^2} \sum_{k=1}^n \left[V_k^2 - \mathbb{E}(V_k^2 \,|\, \mathcal{F}_{k-1}) \right] + \frac{1}{1-(\alpha-\beta)^2} \langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \, \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle \sum_{k=1}^n U_{k-1} \\
&\quad - \frac{(\alpha-\beta)^2}{1-(\alpha-\beta)^2} V_n^2 + \mathrm{O}(n) + \text{constant} \times \sum_{k=1}^n V_{k-1}.
\end{aligned}$$

Using (B.7) with $\ell = 8$, $i = 0$ and $j = 2$ we obtain

$$\frac{1}{n^2} \sum_{k=1}^n \left[V_k^2 - \mathbb{E}(V_k^2 \,|\, \mathcal{F}_{k-1}) \right] \stackrel{\mathbb{P}}{\longrightarrow} 0 \qquad \text{as } n \to \infty.$$

By Corollary B.7, we obtain $\mathbb{E}(V_n^2) = \mathrm{O}(n)$, and hence $n^{-2} V_n^2 \stackrel{\mathbb{P}}{\longrightarrow} 0$. Moreover, $n^{-2} \sum_{k=1}^n V_{k-1} \stackrel{\mathbb{P}}{\longrightarrow} 0$ as $n \to \infty$ follows by (B.5) with the choices $\ell = 8$, $i = 0$, $j = 1$. Consequently, by (A.5), we obtain the statement, since $1 - (\alpha-\beta)^2 = 4\alpha\beta$. $\square$

A.3 Lemma. *If* $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle = 0$ *then*

$$n^{-1} \sum_{k=1}^n V_k^2 \stackrel{\text{a.s.}}{\longrightarrow} \mathbb{E}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_1 \rangle^2) \qquad \text{as } n \to \infty,$$

and $\mathbb{E}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_1 \rangle^2) = 0$ *if and only if* $X_{k,1} \stackrel{\text{a.s.}}{=} X_{k,2}$ *for all* $k \in \mathbb{N}$.

Proof. By Remark 3.3, $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle = 0$ implies

$$V_k = \sum_{j=1}^{X_{k-1,1}} (\xi_{k,j,1,1} - \xi_{k,j,1,2}) + \sum_{j=1}^{X_{k-1,2}} (\xi_{k,j,2,1} - \xi_{k,j,2,2}) + (\varepsilon_{k,1} - \varepsilon_{k,2}) \stackrel{\text{a.s.}}{=} \varepsilon_{k,1} - \varepsilon_{k,2} = \langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_1 \rangle$$

for all $k \in \mathbb{N}$, hence the convergence follows from the strong law of large numbers. Clearly $\mathbb{E}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_1 \rangle^2) = 0$ is equivalent to $\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_1 \rangle = \varepsilon_{1,1} - \varepsilon_{1,2} \stackrel{\text{a.s.}}{=} 0$, and hence it is equivalent to $X_{k,1} - X_{k,2} \stackrel{\text{a.s.}}{=} 0$ for all $k \in \mathbb{N}$. $\square$

Now we can prove existence and uniqueness of CLS estimators of the offspring means and of the criticality parameter.

A.4 Proposition. *We have* $\lim_{n\to\infty} \mathbb{P}((\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n) \in H_n) = 1$, *where* H_n *is defined in* (3.2), *and hence the probability of the existence of a unique CLS estimator* $\widehat{\varrho}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$ *converges to 1 as* $n \to \infty$, *and this CLS estimator has the form given in* (3.1) *whenever the sample* $(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$ *belongs to the set* H_n.

If $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle > 0$, *or if* $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}} \rangle = 0$ *and* $\mathbb{E}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon} \rangle^2) > 0$, *then* $\lim_{n\to\infty} \mathbb{P}((\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n) \in \widetilde{H}_n) = 1$, *where* $\widetilde{H}_n$ *is defined in* (3.4), *and hence the probability of the existence of unique CLS estimators* $\widehat{\delta}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$ *and* $(\widehat{\alpha}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n), \widehat{\beta}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n))$ *converges to 1 as* $n \to \infty$. *The CLS estimator* $\widehat{\delta}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$ *has the form given in* (3.5) *whenever the sample* $(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$ *belongs to the set* $\widetilde{H}_n$. *The CLS estimator* $(\widehat{\alpha}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n), \widehat{\beta}_n(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n))$ *has the form given in* (3.3) *whenever the sample* $(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$ *belongs to the set* $H_n \cap \widetilde{H}_n$.

Proof. Recall convergence $\boldsymbol{\mathcal{X}}^{(n)} \xrightarrow{\mathcal{D}} \boldsymbol{\mathcal{X}} = \frac{1}{2}\mathcal{Y}\boldsymbol{1}$ from (3.11). First we show

$$\text{(A.6)} \qquad \frac{1}{n^3} \sum_{k=1}^{n} (X_{k-1,1}^2 + X_{k-1,2}^2) = \frac{1}{n^3} \sum_{k=1}^{n} \|\boldsymbol{X}_{k-1}\|^2 \xrightarrow{\mathcal{D}} \int_0^1 \|\boldsymbol{\mathcal{X}}_t\|^2 \, \mathrm{d}t = \frac{1}{2} \int_0^1 \mathcal{Y}_t^2 \, \mathrm{d}t$$

as $n \to \infty$. Let us apply Lemmas C.2 and C.3 with the special choices $d := 2$, $p := q := 1$, $h : \mathbb{R}^2 \to \mathbb{R}$, $h(\boldsymbol{x}) := \boldsymbol{x}$, $\boldsymbol{x} \in \mathbb{R}^2$, $K : [0,1] \times \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$,

$$K(s, \boldsymbol{x}, \boldsymbol{y}) := \|\boldsymbol{x}\|^2, \qquad (s, \boldsymbol{x}, \boldsymbol{y}) \in [0,1] \times \mathbb{R}^2 \times \mathbb{R}^2,$$

and $\boldsymbol{\mathcal{U}} := \boldsymbol{\mathcal{X}}$, $\boldsymbol{\mathcal{U}}^{(n)} := \boldsymbol{\mathcal{X}}^{(n)}$, $n \in \mathbb{N}$. Then

$$\begin{aligned}
|K(s, \boldsymbol{x}, \boldsymbol{y}) - K(t, \boldsymbol{u}, \boldsymbol{v})| &= \big| \|\boldsymbol{x}\|^2 - \|\boldsymbol{u}\|^2 \big| \leqslant (\|\boldsymbol{x}\| + \|\boldsymbol{u}\|) \big| \|\boldsymbol{x}\| - \|\boldsymbol{u}\| \big| \\
&\leqslant 2R\Big(|t-s| + \big| \|\boldsymbol{x}\| - \|\boldsymbol{u}\| \big| \Big) \leqslant 2R\big(|t-s| + \|(\boldsymbol{x}, \boldsymbol{y}) - (\boldsymbol{u}, \boldsymbol{v})\| \big)
\end{aligned}$$

for all $s, t \in [0,1]$ and $(\boldsymbol{x}, \boldsymbol{y}), (\boldsymbol{u}, \boldsymbol{v}) \in \mathbb{R}^2 \times \mathbb{R}^2$ with $\|(\boldsymbol{x}, \boldsymbol{y})\| \leqslant R$ and $\|(\boldsymbol{u}, \boldsymbol{v})\| \leqslant R$, where $R > 0$. Further, using the definition of Φ and Φ_n, $n \in \mathbb{N}$, given in Lemma C.3,

$$\Phi_n(\boldsymbol{\mathcal{X}}^{(n)}) = \left(\boldsymbol{\mathcal{X}}_1^{(n)}, \frac{1}{n} \sum_{k=1}^{n} \|\boldsymbol{\mathcal{X}}_{k/n}^{(n)}\|^2 \right) = \left(\frac{1}{n} \boldsymbol{X}_n, \frac{1}{n^3} \sum_{k=1}^{n} \|\boldsymbol{X}_k\|^2 \right),$$

$$\Phi(\boldsymbol{\mathcal{X}}) = \left(\boldsymbol{\mathcal{X}}_1, \int_0^1 \|\boldsymbol{\mathcal{X}}_u\|^2 \, \mathrm{d}u \right).$$

Since the process $(\boldsymbol{\mathcal{X}}_t)_{t\in\mathbb{R}_+}$ admits continuous paths with probability one, Lemma C.2 (with the choice $C := \mathsf{C}(\mathbb{R}_+,\mathbb{R})$) and Lemma C.3 yield (A.6). Since $\boldsymbol{m}_{\boldsymbol{\varepsilon}} \neq \boldsymbol{0}$, by the SDE (3.6), we have $\mathbb{P}\big(\mathcal{Y}_t = 0,\, t \in [0,1]\big) = 0$, which implies that $\mathbb{P}\big(\int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t > 0\big) = 1$. Consequently, the distribution function of $\int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t$ is continuous at 0, and hence, by (A.6),

$$\mathbb{P}\left(\sum_{k=1}^n \langle \boldsymbol{1}, \boldsymbol{X}_{k-1}\rangle^2 > 0\right) = \mathbb{P}\left(\frac{1}{n^3}\sum_{k=1}^n \|\boldsymbol{X}_{k-1}\|^2 > 0\right) \to \mathbb{P}\left(\frac{1}{2}\int_0^1 \mathcal{Y}_t^2 \,\mathrm{d}t > 0\right) = 1$$

as $n \to \infty$.

Now suppose that $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle > 0$ holds. In a similar way, using Lemma A.2, convergence (3.11), and Lemmas C.2 and C.3, one can show

$$\frac{1}{n^2}\sum_{k=1}^n \langle \widetilde{\boldsymbol{u}}, \boldsymbol{X}_{k-1}\rangle^2 \stackrel{\mathcal{D}}{\longrightarrow} \frac{\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle}{4\alpha\beta}\int_0^1 \mathcal{Y}_t \,\mathrm{d}t \qquad \text{as } n \to \infty,$$

implying

$$\mathbb{P}\left(\sum_{k=1}^n \langle \widetilde{\boldsymbol{u}}, \boldsymbol{X}_{k-1}\rangle^2 > 0\right) \to \mathbb{P}\left(\int_0^1 \mathcal{Y}_t \,\mathrm{d}t > 0\right) = 1,$$

hence we obtain the statement under the assumption $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle > 0$.

Next we suppose that $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle = 0$ and $\mathbb{E}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}\rangle^2) > 0$ hold. Then

$$\mathbb{P}\left(\sum_{k=1}^n \langle \widetilde{\boldsymbol{u}}, \boldsymbol{X}_{k-1}\rangle^2 > 0\right) = \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^n V_{k-1}^2 > 0\right) \to 1,$$

since Lemma A.3 yields $n^{-1}\sum_{k=1}^n V_{k-1}^2 \stackrel{\mathbb{P}}{\longrightarrow} \mathbb{E}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}_1\rangle^2) > 0$, and hence we conclude the statement under the assumptions $\langle \overline{\boldsymbol{V}}_{\boldsymbol{\xi}} \widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{u}}\rangle = 0$ and $\mathbb{E}(\langle \widetilde{\boldsymbol{u}}, \boldsymbol{\varepsilon}\rangle^2) > 0$. □

B Estimations of moments

In the proof of Theorem 3.1, good bounds for moments of the random vectors and variables $(\boldsymbol{M}_k)_{k\in\mathbb{Z}_+}$, $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$, $(U_k)_{k\in\mathbb{Z}_+}$ and $(V_k)_{k\in\mathbb{Z}_+}$ are extensively used. First note that, for all $k \in \mathbb{N}$, $\mathbb{E}(\boldsymbol{M}_k \mid \mathcal{F}_{k-1}) = \boldsymbol{0}$ and $\mathbb{E}(\boldsymbol{M}_k) = \boldsymbol{0}$, since $\boldsymbol{M}_k = \boldsymbol{X}_k - \mathbb{E}(\boldsymbol{X}_k \mid \mathcal{F}_{k-1})$.

B.1 Lemma. *Let $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ be a 2-type doubly symmetric Galton–Watson process with immigration and with $\boldsymbol{X}_0 = \boldsymbol{0}$. If $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^2) < \infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^2) < \infty$ and $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^2) < \infty$ then*

$$\text{(B.1)} \qquad \mathbb{E}(\boldsymbol{M}_k \boldsymbol{M}_k^\top \mid \mathcal{F}_{k-1}) = X_{k-1,1}\boldsymbol{V}_{\boldsymbol{\xi}_1} + X_{k-1,2}\boldsymbol{V}_{\boldsymbol{\xi}_2} + \boldsymbol{V}_{\boldsymbol{\varepsilon}}, \qquad k \in \mathbb{N}.$$

If $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^3) < \infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^3) < \infty$ and $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^3) < \infty$ then, for all $k \in \mathbb{N}$,

$$\text{(B.2)} \qquad \begin{aligned} \mathbb{E}(\boldsymbol{M}_k^{\otimes 3} \mid \mathcal{F}_{k-1}) &= X_{k-1,1}\,\mathbb{E}[(\boldsymbol{\xi}_{1,1,1} - \mathbb{E}(\boldsymbol{\xi}_{1,1,1})^{\otimes 3}] \\ &\quad + X_{k-1,2}\,\mathbb{E}[(\boldsymbol{\xi}_{1,1,2} - \mathbb{E}(\boldsymbol{\xi}_{1,1,2})^{\otimes 3}] + \mathbb{E}[(\boldsymbol{\varepsilon}_1 - \mathbb{E}(\boldsymbol{\varepsilon}_1)^{\otimes 3}]. \end{aligned}$$

Proof. By (2.1) and (4.1),

$$\boldsymbol{M}_k = \sum_{j=1}^{X_{k-1,1}} \big(\boldsymbol{\xi}_{k,j,1} - \mathbb{E}(\boldsymbol{\xi}_{k,j,1})\big) + \sum_{j=1}^{X_{k-1,2}} \big(\boldsymbol{\xi}_{k,j,2} - \mathbb{E}(\boldsymbol{\xi}_{k,j,2})\big) + \big(\boldsymbol{\varepsilon}_k - \mathbb{E}(\boldsymbol{\varepsilon}_k)\big), \qquad k \in \mathbb{N}. \tag{B.3}$$

For all $k \in \mathbb{N}$, the random vectors $\big\{\boldsymbol{\xi}_{k,j,1} - \mathbb{E}(\boldsymbol{\xi}_{k,j,1}),\, \boldsymbol{\xi}_{k,j.2} - \mathbb{E}(\boldsymbol{\xi}_{k,j,2}),\, \boldsymbol{\varepsilon}_k - \mathbb{E}(\boldsymbol{\varepsilon}_k) : j \in \mathbb{N}\big\}$ are independent of each other, independent of $\mathcal{F}_{k-1}$, and have zero mean vector, thus we conclude (B.1) and (B.2). $\Box$

B.2 Lemma. *Let $(\boldsymbol{\zeta}_k)_{k\in\mathbb{N}}$ be independent and identically distributed random vectors with values in $\mathbb{R}^d$ such that $\mathbb{E}(\|\boldsymbol{\zeta}_1\|^\ell) < \infty$ with some $\ell \in \mathbb{N}$.*

(i) *Then there exists $\boldsymbol{Q} = (Q_1, \dots, Q_{d^\ell}) : \mathbb{R} \to \mathbb{R}^{d^\ell}$, where $Q_1, \dots, Q_{d^\ell}$ are polynomials having degree at most $\ell - 1$ such that*

$$\mathbb{E}\big((\boldsymbol{\zeta}_1 + \dots + \boldsymbol{\zeta}_N)^{\otimes\ell}\big) = N^\ell \big[\mathbb{E}(\boldsymbol{\zeta}_1)\big]^{\otimes\ell} + \boldsymbol{Q}(N), \qquad N \in \mathbb{N}, \quad N \geqslant \ell.$$

(ii) *If $\mathbb{E}(\boldsymbol{\zeta}_1) = \boldsymbol{0}$ then there exists $\boldsymbol{R} = (R_1, \dots, R_{d^\ell}) : \mathbb{R} \to \mathbb{R}^{d^\ell}$, where $R_1, \dots, R_{d^\ell}$ are polynomials having degree at most $\lfloor \ell/2 \rfloor$ such that*

$$\mathbb{E}\big((\boldsymbol{\zeta}_1 + \dots + \boldsymbol{\zeta}_N)^{\otimes\ell}\big) = \boldsymbol{R}(N), \qquad N \in \mathbb{N}, \quad N \geqslant \ell.$$

The coefficients of the polynomials $\boldsymbol{Q}$ and $\boldsymbol{R}$ depend on the moments $\mathbb{E}(\boldsymbol{\zeta}_{i_1} \otimes \dots \otimes \boldsymbol{\zeta}_{i_\ell})$, $i_1, \dots, i_\ell \in \{1, \dots, N\}$.

Proof. (i) We have

$$\mathbb{E}\big((\boldsymbol{\zeta}_1 + \dots + \boldsymbol{\zeta}_N)^{\otimes\ell}\big) = \sum_{\substack{s\in\{1,\dots,\ell\},\, k_1,\dots,k_s\in\mathbb{Z}_+,\\ k_1+2k_2+\dots+sk_s=\ell,\, k_s\neq 0}} \binom{N}{k_1}\binom{N-k_1}{k_2}\cdots\binom{N-k_1-\dots-k_{s-1}}{k_s} \times \sum_{(i_1,\dots,i_\ell)\in P^{(N,\ell)}_{k_1,\dots,k_s}} \mathbb{E}(\boldsymbol{\zeta}_{i_1} \otimes \dots \otimes \boldsymbol{\zeta}_{i_\ell}),$$

where the set $P^{(N,\ell)}_{k_1,\dots,k_s}$ consists of permutations of all the multisets containing pairwise different elements $j_{k_1}, \dots, j_{k_s}$ of the set $\{1, \dots, N\}$ with multiplicities $k_1, \dots, k_s$, respectively. Since

$$\binom{N}{k_1}\binom{N-k_1}{k_2}\cdots\binom{N-k_1-\dots-k_{s-1}}{k_s} = \frac{N(N-1)\cdots(N-k_1-k_2-\dots-k_s+1)}{k_1!k_2!\cdots k_s!}$$

is a polynomial of the variable N having degree $k_1 + \dots + k_s \leqslant \ell$, there exists $\boldsymbol{P} = (P_1, \dots, P_{d^\ell}) : \mathbb{R} \to \mathbb{R}^{d^\ell}$, where $P_1, \dots, P_{d^\ell}$ are polynomials having degree at most ℓ such that $\mathbb{E}\big((\boldsymbol{\zeta}_1 + \dots + \boldsymbol{\zeta}_N)^{\otimes\ell}\big) = \boldsymbol{P}(N)$. A term of degree ℓ can occur only in case $k_1 + \dots + k_s = \ell$, when $k_1 + 2k_2 + \dots + sk_s = \ell$ implies $s = 1$ and $k_1 = \ell$, thus the corresponding term of degree ℓ is $N(N-1)\cdots(N-\ell+1)\big[\mathbb{E}(\boldsymbol{\zeta}_1)\big]^{\otimes\ell}$, hence we obtain the statement.

(ii) Using the same decomposition, we have

$$\mathbb{E}\big((\boldsymbol{\zeta}_1+\cdots+\boldsymbol{\zeta}_N)^{\otimes\ell}\big) = \sum_{\substack{s\in\{2,\ldots,\ell\},\, k_2,\ldots,k_s\in\mathbb{Z}_+,\\ 2k_2+3k_3+\cdots+sk_s=\ell,\, k_s\neq 0}} \binom{N}{k_2}\binom{N-k_2}{k_3}\cdots\binom{N-k_2-\cdots-k_{s-1}}{k_s}$$

$$\times \sum_{(i_1,\ldots,i_\ell)\in P^{(N,\ell)}_{0,k_2,\ldots,k_s}} \mathbb{E}(\boldsymbol{\zeta}_{i_1}\otimes\cdots\otimes\boldsymbol{\zeta}_{i_\ell}).$$

Here

$$\binom{N}{k_2}\binom{N-k_2}{k_3}\cdots\binom{N-k_2-\cdots-k_{s-1}}{k_s} = \frac{N(N-1)\cdots(N-k_2-k_3-\cdots-k_s+1)}{k_2!k_3!\cdots k_s!}$$

is a polynomial of the variable N having degree $k_2+\cdots+k_s$. Since

$$\ell = 2k_2+3k_3+\cdots+sk_s \geqslant 2(k_2+k_3+\cdots+k_s),$$

we have $k_2+\cdots+k_s\leqslant \ell/2$ yielding part (ii). □

B.3 Remark. In what follows, using the proof of Lemma B.4, we give a bit more explicit form of the polynomial R_ℓ in part (ii) of Lemma B.4 for the special cases $\ell=1,2,3,4,5,6$.

$$\mathbb{E}(\boldsymbol{\zeta}_1+\cdots+\boldsymbol{\zeta}_N) = \mathbf{0}$$

$$\mathbb{E}((\boldsymbol{\zeta}_1+\cdots+\boldsymbol{\zeta}_N)^{\otimes 2}) = N\,\mathbb{E}(\boldsymbol{\zeta}_1^{\otimes 2}).$$

$$\mathbb{E}((\boldsymbol{\zeta}_1+\cdots+\boldsymbol{\zeta}_N)^{\otimes 3}) = N\,\mathbb{E}(\boldsymbol{\zeta}_1^{\otimes 3}).$$

$$\mathbb{E}((\boldsymbol{\zeta}_1+\cdots+\boldsymbol{\zeta}_N)^{\otimes 4}) = N\,\mathbb{E}(\boldsymbol{\zeta}_1^{\otimes 4}) + \frac{N(N-1)}{2!}\sum_{(i_1,i_2,i_3,i_4)\in P^{(N,4)}_{0,2}} \mathbb{E}(\boldsymbol{\zeta}_{i_1}\otimes\boldsymbol{\zeta}_{i_2}\otimes\boldsymbol{\zeta}_{i_3}\otimes\boldsymbol{\zeta}_{i_4}).$$

$$\mathbb{E}((\boldsymbol{\zeta}_1+\cdots+\boldsymbol{\zeta}_N)^{\otimes 5}) = N\,\mathbb{E}(\boldsymbol{\zeta}_1^{\otimes 5}) + N(N-1)\sum_{(i_1,i_2,i_3,i_4,i_5)\in P^{(N,5)}_{0,1,1}} \mathbb{E}(\boldsymbol{\zeta}_{i_1}\otimes\boldsymbol{\zeta}_{i_2}\otimes\boldsymbol{\zeta}_{i_3}\otimes\boldsymbol{\zeta}_{i_4}\otimes\boldsymbol{\zeta}_{i_5}).$$

$$\begin{aligned}
&\mathbb{E}((\boldsymbol{\zeta}_1+\cdots+\boldsymbol{\zeta}_N)^{\otimes 6})\\
&\quad= N\,\mathbb{E}(\boldsymbol{\zeta}_1^{\otimes 6}) + N(N-1)\sum_{(i_1,i_2,i_3,i_4,i_5,i_6)\in P^{(N,6)}_{0,1,0,1}} \mathbb{E}(\boldsymbol{\zeta}_{i_1}\otimes\boldsymbol{\zeta}_{i_2}\otimes\boldsymbol{\zeta}_{i_3}\otimes\boldsymbol{\zeta}_{i_4}\otimes\boldsymbol{\zeta}_{i_5}\otimes\boldsymbol{\zeta}_{i_6})\\
&\qquad+\frac{N(N-1)}{2!}\sum_{(i_1,i_2,i_3,i_4,i_5,i_6)\in P^{(N,6)}_{0,0,2}} \mathbb{E}(\boldsymbol{\zeta}_{i_1}\otimes\boldsymbol{\zeta}_{i_2}\otimes\boldsymbol{\zeta}_{i_3}\otimes\boldsymbol{\zeta}_{i_4}\otimes\boldsymbol{\zeta}_{i_5}\otimes\boldsymbol{\zeta}_{i_6})\\
&\qquad+\frac{N(N-1)(N-2)}{3!}\sum_{(i_1,i_2,i_3,i_4,i_5,i_6)\in P^{(N,6)}_{0,3}} \mathbb{E}(\boldsymbol{\zeta}_{i_1}\otimes\boldsymbol{\zeta}_{i_2}\otimes\boldsymbol{\zeta}_{i_3}\otimes\boldsymbol{\zeta}_{i_4}\otimes\boldsymbol{\zeta}_{i_5}\otimes\boldsymbol{\zeta}_{i_6}).
\end{aligned}$$

□

Lemma B.2 can be generalized in the following way.

B.4 Lemma. *For each* $i \in \mathbb{N}$, *let* $(\boldsymbol{\zeta}_{i,k})_{k\in\mathbb{N}}$ *be independent and identically distributed random vectors with values in* $\mathbb{R}^d$ *such that* $\mathbb{E}(\|\boldsymbol{\zeta}_{i,1}\|^\ell) < \infty$ *with some* $\ell \in \mathbb{N}$. *Let* $j_1, \dots, j_\ell \in \mathbb{N}$.

(i) *Then there exists* $\boldsymbol{Q} = (Q_1, \dots, Q_{d^\ell}) : \mathbb{R}^\ell \to \mathbb{R}^{d^\ell}$, *where* $Q_1, \dots, Q_{d^\ell}$ *are polynomials of* ℓ *variables having degree at most* $\ell - 1$ *such that*

$$\begin{aligned}&\mathbb{E}\big((\boldsymbol{\zeta}_{j_1,1} + \dots + \boldsymbol{\zeta}_{j_1,N_1}) \otimes \dots \otimes (\boldsymbol{\zeta}_{j_\ell,1} + \dots + \boldsymbol{\zeta}_{j_\ell,N_\ell})\big)\\ &\qquad = N_1 \dots N_\ell \, \mathbb{E}(\boldsymbol{\zeta}_{j_1,1}) \otimes \dots \otimes \mathbb{E}(\boldsymbol{\zeta}_{j_\ell,1}) + \boldsymbol{Q}(N_1, \dots, N_\ell)\end{aligned}$$

for $N_1, \dots, N_\ell \in \mathbb{N}$ *with* $N_1 \geqslant \ell, \dots, N_\ell \geqslant \ell$.

(ii) *If* $\mathbb{E}(\boldsymbol{\zeta}_{j_1,1}) = \dots = \mathbb{E}(\boldsymbol{\zeta}_{j_\ell,1}) = \boldsymbol{0}$ *then there exists* $\boldsymbol{R} = (R_1, \dots, R_{d^\ell}) : \mathbb{R}^\ell \to \mathbb{R}^{d^\ell}$, *where* $R_1, \dots, R_{d^\ell}$ *are polynomials of* ℓ *variables having degree at most* $\lfloor \ell/2 \rfloor$ *such that*

$$\mathbb{E}\big((\boldsymbol{\zeta}_{j_1,1} + \dots + \boldsymbol{\zeta}_{j_1,N_1}) \otimes \dots \otimes (\boldsymbol{\zeta}_{j_\ell,1} + \dots + \boldsymbol{\zeta}_{j_\ell,N_\ell})\big) = \boldsymbol{R}(N_1, \dots, N_\ell)$$

for $N_1, \dots, N_\ell \in \mathbb{N}$ *with* $N_1 \geqslant \ell, \dots, N_\ell \geqslant \ell$.

The coefficients of the polynomials $\boldsymbol{Q}$ *and* $\boldsymbol{R}$ *depend on the moments* $\mathbb{E}(\boldsymbol{\zeta}_{j_1,i_1} \otimes \dots \otimes \boldsymbol{\zeta}_{j_\ell,i_\ell})$, $i_1 \in \{1, \dots, N_1\}, \dots, i_\ell \in \{1, \dots, N_\ell\}$.

B.5 Lemma. *If* $(\alpha, \beta) \in [0, 1]$ *with* $\alpha + \beta = 1$, *then the matrix* $\boldsymbol{m}_{\boldsymbol{\xi}}$ *defined in* (2.4) *has eigenvalues* 1 *and* $\alpha - \beta$, *and the powers of* $\boldsymbol{m}_{\boldsymbol{\xi}}$ *take the form*

$$\boldsymbol{m}_{\boldsymbol{\xi}}^j = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{1}{2}(\alpha - \beta)^j \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \qquad j \in \mathbb{Z}_+.$$

Consequently, $\|\boldsymbol{m}_{\boldsymbol{\xi}}^j\| = \mathrm{O}(1)$, *i.e.,* $\sup_{j\in\mathbb{N}} \|\boldsymbol{m}_{\boldsymbol{\xi}}^j\| < \infty$.

Proof. The formula for the powers of $\boldsymbol{m}_{\boldsymbol{\xi}}$ follows by the so-called Putzer's spectral formula, see, e.g., Putzer [12]. □

B.6 Lemma. *Let* $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ *be a 2-type doubly symmetric Galton–Watson process with immigration with offspring means* $(\alpha, \beta) \in [0, 1]$ *such that* $\alpha + \beta = 1$ *(hence it is critical). Suppose* $\boldsymbol{X}_0 = \boldsymbol{0}$, *and* $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^\ell) < \infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^\ell) < \infty$, $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^\ell) < \infty$ *with some* $\ell \in \mathbb{N}$. *Then* $\mathbb{E}(\|\boldsymbol{X}_k\|^\ell) = \mathrm{O}(k^\ell)$, *i.e.,* $\sup_{k\in\mathbb{N}} k^{-\ell}\,\mathbb{E}(\|\boldsymbol{X}_k\|^\ell) < \infty$.

Proof. The statement is clearly equivalent with $\mathbb{E}\big(P(X_{k,1}, X_{k,2})\big) \leqslant c_P\, k^\ell$, $k \in \mathbb{N}$, for all polynomials P of two variables having degree at most ℓ, where c_P depends only on P.

If $\ell = 1$ then (2.3) and Lemma B.5 imply

$$\mathbb{E}(\boldsymbol{X}_k) = \sum_{j=0}^{k-1} \boldsymbol{m}_{\boldsymbol{\xi}}^j \, \boldsymbol{m}_{\boldsymbol{\varepsilon}} = \left(\frac{k}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{1 - (\alpha - \beta)^k}{4\beta}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\right)\boldsymbol{m}_{\boldsymbol{\varepsilon}}, \quad k \in \mathbb{N},$$

which yields the statement.

By (2.1),

$$\boldsymbol{X}_k^{\otimes 2} = \left(\sum_{j=1}^{X_{k-1,1}} \boldsymbol{\xi}_{k,j,1}\right)^{\otimes 2} + \left(\sum_{j=1}^{X_{k-1,2}} \boldsymbol{\xi}_{k,j,2}\right)^{\otimes 2} + \boldsymbol{\varepsilon}_k^{\otimes 2} + \left(\sum_{j=1}^{X_{k-1,1}} \boldsymbol{\xi}_{k,j,1}\right) \otimes \left(\sum_{j=1}^{X_{k-1,2}} \boldsymbol{\xi}_{k,j,2}\right)$$

$$\text{(B.4)} \qquad + \left(\sum_{j=1}^{X_{k-1,2}} \boldsymbol{\xi}_{k,j,2}\right) \otimes \left(\sum_{j=1}^{X_{k-1,1}} \boldsymbol{\xi}_{k,j,1}\right) + \left(\sum_{j=1}^{X_{k-1,1}} \boldsymbol{\xi}_{k,j,1}\right) \otimes \boldsymbol{\varepsilon}_k + \boldsymbol{\varepsilon}_k \otimes \left(\sum_{j=1}^{X_{k-1,1}} \boldsymbol{\xi}_{k,j,1}\right)$$

$$+ \left(\sum_{j=1}^{X_{k-1,2}} \boldsymbol{\xi}_{k,j,1}\right) \otimes \boldsymbol{\varepsilon}_k + \boldsymbol{\varepsilon}_k \otimes \left(\sum_{j=1}^{X_{k-1,2}} \boldsymbol{\xi}_{k,j,1}\right).$$

Since for all $k \in \mathbb{N}$, the random variables $\{\boldsymbol{\xi}_{k,j,1}, \boldsymbol{\xi}_{k,j,2}, \boldsymbol{\varepsilon}_k : j \in \mathbb{N}\}$ are independent of each other and of the σ-algebra $\mathcal{F}_{k-1}$, we have

$$\mathbb{E}(\boldsymbol{X}_k^{\otimes 2} \,|\, \mathcal{F}_{k-1}) = \mathbb{E}\left(\left(\sum_{j=1}^{M} \boldsymbol{\xi}_{k,j,1}\right)^{\otimes 2}\right)\Bigg|_{M=X_{k-1,1}} + \mathbb{E}\left(\left(\sum_{j=1}^{N} \boldsymbol{\xi}_{k,j,2}\right)^{\otimes 2}\right)\Bigg|_{N=X_{k-1,2}} + \mathbb{E}(\boldsymbol{\varepsilon}_k^{\otimes 2})$$

$$+ \mathbb{E}\left(\sum_{j=1}^{M} \boldsymbol{\xi}_{k,j,1}\right) \otimes \mathbb{E}\left(\sum_{j=1}^{N} \boldsymbol{\xi}_{k,j,2}\right)\Bigg|_{\substack{M=X_{k-1,1}\\ N=X_{k-1,2}}} + \mathbb{E}\left(\sum_{j=1}^{N} \boldsymbol{\xi}_{k,j,2}\right) \otimes \mathbb{E}\left(\sum_{j=1}^{M} \boldsymbol{\xi}_{k,j,1}\right)\Bigg|_{\substack{M=X_{k-1,1}\\ N=X_{k-1,2}}}$$

$$+ \mathbb{E}\left(\sum_{j=1}^{M} \boldsymbol{\xi}_{k,j,1}\right)\Bigg|_{M=X_{k-1,1}} \otimes \mathbb{E}(\boldsymbol{\varepsilon}_k) + \mathbb{E}(\boldsymbol{\varepsilon}_n) \otimes \mathbb{E}\left(\sum_{j=1}^{M} \boldsymbol{\xi}_{k,j,1}\right)\Bigg|_{M=X_{k-1,1}}$$

$$+ \mathbb{E}\left(\sum_{j=1}^{N} \boldsymbol{\xi}_{k,j,2}\right)\Bigg|_{N=X_{k-1,2}} \otimes \mathbb{E}(\boldsymbol{\varepsilon}_k) + \mathbb{E}(\boldsymbol{\varepsilon}_k) \otimes \mathbb{E}\left(\sum_{j=1}^{N} \boldsymbol{\xi}_{k,j,2}\right)\Bigg|_{N=X_{k-1,2}}.$$

Using part (i) of Lemma B.4 and separating the terms having degree 2 and less than 2, we have

$$\begin{aligned}
&\mathbb{E}(\boldsymbol{X}_k^{\otimes 2} \,|\, \mathcal{F}_{k-1}) \\
&= X_{k-1,1}^2 \boldsymbol{m}_{\boldsymbol{\xi}_1}^{\otimes 2} + X_{k-1,2}^2 \boldsymbol{m}_{\boldsymbol{\xi}_2}^{\otimes 2} + X_{k-1,1}X_{k-1,2}(\boldsymbol{m}_{\boldsymbol{\xi}_1} \otimes \boldsymbol{m}_{\boldsymbol{\xi}_2} + \boldsymbol{m}_{\boldsymbol{\xi}_2} \otimes \boldsymbol{m}_{\boldsymbol{\xi}_1}) + \boldsymbol{Q}_2(X_{k-1,1}, X_{k-1,2}) \\
&= (X_{k-1,1}\boldsymbol{m}_{\boldsymbol{\xi}_1} + X_{k-1,2}\boldsymbol{m}_{\boldsymbol{\xi}_2})^{\otimes 2} + \boldsymbol{Q}_2(X_{k-1,1}, X_{k-1,2}) = (\boldsymbol{m}_{\boldsymbol{\xi}}\boldsymbol{X}_{k-1})^{\otimes 2} + \boldsymbol{Q}_2(X_{k-1,1}, X_{k-1,2}) \\
&= \boldsymbol{m}_{\boldsymbol{\xi}}^{\otimes 2} \boldsymbol{X}_{k-1}^{\otimes 2} + \boldsymbol{Q}_2(X_{k-1,1}, X_{k-1,2}),
\end{aligned}$$

where $\boldsymbol{Q}_2 = (Q_{2,1}, Q_{2,2}, Q_{2,3}, Q_{2,4}) : \mathbb{R}^2 \to \mathbb{R}^4$, and $Q_{2,1}$, $Q_{2,2}$, $Q_{2,3}$ and $Q_{2,4}$ are polynomials of two variables having degree at most 1. Hence

$$\mathbb{E}(\boldsymbol{X}_k^{\otimes 2}) = \boldsymbol{m}_{\boldsymbol{\xi}}^{\otimes 2}\, \mathbb{E}(\boldsymbol{X}_{k-1}^{\otimes 2}) + \mathbb{E}[\boldsymbol{Q}_2(X_{k-1,1}, X_{k-1,2})].$$

In a similar way,

$$\mathbb{E}(\boldsymbol{X}_k^{\otimes \ell}) = \boldsymbol{m}_{\boldsymbol{\xi}}^{\otimes \ell}\, \mathbb{E}(\boldsymbol{X}_{k-1}^{\otimes \ell}) + \mathbb{E}[\boldsymbol{Q}_\ell(X_{k-1,1}, X_{k-1,2})],$$

where $\boldsymbol{Q}_\ell = (Q_{\ell,1}, \ldots, Q_{\ell,2^\ell}) : \mathbb{R}^2 \to \mathbb{R}^{2^\ell}$, and $Q_{\ell,1}, \ldots, Q_{\ell,2^\ell}$ are polynomials of two variables having degree at most $\ell - 1$, implying

$$\begin{aligned}
\mathbb{E}(\boldsymbol{X}_k^{\otimes \ell}) &= \sum_{j=1}^{k} (\boldsymbol{m}_{\boldsymbol{\xi}}^{\otimes \ell})^{k-j}\, \mathbb{E}[\boldsymbol{Q}_\ell(X_{j-1,1}, X_{j-1,2})] = \sum_{j=0}^{k-1} (\boldsymbol{m}_{\boldsymbol{\xi}}^{\otimes \ell})^{j}\, \mathbb{E}[\boldsymbol{Q}_\ell(X_{k-j-1,1}, X_{k-j-1,2})] \\
&= \sum_{j=0}^{k-1} (\boldsymbol{m}_{\boldsymbol{\xi}}^{j})^{\otimes \ell}\, \mathbb{E}[\boldsymbol{Q}_\ell(X_{k-j-1,1}, X_{k-j-1,2})].
\end{aligned}$$

Let us suppose now that the statement holds for $1, \ldots, \ell - 1$. Then

$$\mathbb{E}[Q_{\ell,i}(X_{k-j-1,1}, X_{k-j-1,2})] \leqslant c_{Q_{\ell,i}} k^{\ell-1}, \qquad k \in \mathbb{N}, \quad i \in \{1, \ldots, 2^\ell\}.$$

Lemma B.5 clearly implies $\|(\boldsymbol{m}_{\boldsymbol{\xi}}^{j})^{\otimes \ell}\| = \mathrm{O}(1)$, i.e., $\sup_{j\in\mathbb{N}} \|(\boldsymbol{m}_{\boldsymbol{\xi}}^{j})^{\otimes \ell}\| < \infty$, hence we obtain the assertion for ℓ. $\Box$

B.7 Corollary. *Let $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ be a 2-type doubly symmetric Galton–Watson process with immigration with offspring means $(\alpha, \beta) \in (0,1)^2$ such that $\alpha+\beta = 1$ (hence it is critical and positively regular). Suppose $\boldsymbol{X}_0 = \boldsymbol{0}$, and $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^\ell) < \infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^\ell) < \infty$, $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^\ell) < \infty$ with some $\ell \in \mathbb{N}$. Then*

$$\mathbb{E}(\|\boldsymbol{X}_k\|^\ell) = \mathrm{O}(k^\ell), \qquad \mathbb{E}(\boldsymbol{M}_k^{\otimes \ell}) = \mathrm{O}(k^{\lfloor \ell/2 \rfloor}), \qquad \mathbb{E}(U_k^\ell) = \mathrm{O}(k^\ell), \qquad \mathbb{E}(V_k^{2j}) = \mathrm{O}(k^j)$$

for $j \in \mathbb{Z}_+$ with $2j \leqslant \ell$.

Proof. The first statement is just Lemma B.6. Next we turn to prove $\mathbb{E}(\boldsymbol{M}_k^{\otimes \ell}) = \mathrm{O}(k^{\lfloor \ell/2 \rfloor})$. Using (B.3), part (ii) of Lemma B.4, and that the random vectors $\big\{\boldsymbol{\xi}_{k,j,1} - \mathbb{E}(\boldsymbol{\xi}_{k,j,1}), \boldsymbol{\xi}_{k,j.2} - \mathbb{E}(\boldsymbol{\xi}_{k,j,2}), \boldsymbol{\varepsilon}_k - \mathbb{E}(\boldsymbol{\varepsilon}_k) : j \in \mathbb{N}\big\}$ are independent of each other, independent of $\mathcal{F}_{k-1}$, and have zero mean vector, we obtain

$$\mathbb{E}(\boldsymbol{M}_k^{\otimes \ell} \,|\, \mathcal{F}_{k-1}) = \boldsymbol{R}(X_{k-1,1}, X_{k-1,2}),$$

with $\boldsymbol{R} = (R_1, \ldots, R_{2^\ell}) : \mathbb{R}^2 \to \mathbb{R}^{2\ell}$, where $R_1, \ldots, R_{2^\ell}$ are polynomials of two variables having degree at most $\ell/2$. Hence

$$\mathbb{E}(\boldsymbol{M}_k^{\otimes \ell}) = \mathbb{E}(\boldsymbol{R}(X_{k-1,1}, X_{k-1,2})).$$

By Lemma B.6, we conclude $\mathbb{E}(\boldsymbol{M}_k^{\otimes \ell}) = \mathrm{O}(k^{\lfloor \ell/2 \rfloor})$.

Lemma B.6 implies $\mathbb{E}(U_k^\ell) = \mathbb{E}[(X_{k,1} + X_{k,2})^\ell] = \mathrm{O}(k^\ell)$.

Finally, for $j \in \mathbb{Z}_+$ with $2j \leqslant \ell$, we prove $\mathbb{E}(V_k^{2j}) = \mathrm{O}(k^j)$ using induction in k. By the recursion $V_k = (\alpha-\beta)V_{k-1} + \langle \widetilde{\boldsymbol{u}}, \boldsymbol{M}_k + \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle$, $k \in \mathbb{N}$, we have $\mathbb{E}(V_k) = (\alpha-\beta)\,\mathbb{E}(V_{k-1}) + \langle \widetilde{\boldsymbol{u}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle$, $k \in \mathbb{N}$, with initial value $\mathbb{E}(V_0) = 0$, hence

$$\mathbb{E}(V_k) = \langle \widetilde{\boldsymbol{u}}, \boldsymbol{m}_{\boldsymbol{\varepsilon}} \rangle \sum_{i=0}^{k-1} (\alpha - \beta)^i, \qquad k \in \mathbb{N},$$

which yields $\mathbb{E}(|V_k|) = \mathrm{O}(1)$. Indeed, for all $k \in \mathbb{N}$,

$$\left|\sum_{i=0}^{k-1} (\alpha - \beta)^i\right| \leqslant \frac{1}{1 - |\alpha - \beta|}.$$

The rest of the proof can be carried out as in Corollary 9.1 of Barczy et al. [4]. □

The next corollary can be derived exactly as Corollary 9.2 of Barczy et al. [4].

B.8 Corollary. *Let $(\boldsymbol{X}_k)_{k\in\mathbb{Z}_+}$ be a 2-type doubly symmetric Galton–Watson process with immigration with offspring means $(\alpha, \beta) \in (0,1)^2$ such that $\alpha+\beta = 1$ (hence it is critical and positively regular). Suppose $\boldsymbol{X}_0 = \boldsymbol{0}$, and $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,1}\|^\ell) < \infty$, $\mathbb{E}(\|\boldsymbol{\xi}_{1,1,2}\|^\ell) < \infty$, $\mathbb{E}(\|\boldsymbol{\varepsilon}_1\|^\ell) < \infty$ with some $\ell \in \mathbb{N}$. Then*

(i) *for all $i, j \in \mathbb{Z}_+$ with $\max\{i, j\} \leqslant \lfloor \ell/2 \rfloor$, and for all $\kappa > i + \frac{j}{2} + 1$, we have*

$$\text{(B.5)} \qquad n^{-\kappa} \sum_{k=1}^{n} |U_k^i V_k^j| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty,$$

(ii) *for all $i, j \in \mathbb{Z}_+$ with $\max\{i, j\} \leqslant \ell$, for all $T > 0$, and for all $\kappa > i + \frac{j}{2} + \frac{i+j}{\ell}$, we have*

$$\text{(B.6)} \qquad n^{-\kappa} \sup_{t\in[0,T]} |U_{\lfloor nt\rfloor}^i V_{\lfloor nt\rfloor}^j| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty,$$

(iii) *for all $i, j \in \mathbb{Z}_+$ with $\max\{i, j\} \leqslant \lfloor \ell/4 \rfloor$, for all $T > 0$, and for all $\kappa > i + \frac{j}{2} + \frac{1}{2}$, we have*

$$\text{(B.7)} \qquad n^{-\kappa} \sup_{t\in[0,T]} \left| \sum_{k=1}^{\lfloor nt\rfloor} [U_k^i V_k^j - \mathbb{E}(U_k^i V_k^j \mid \mathcal{F}_{k-1})] \right| \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty.$$

B.9 Remark. In the special case $\ell = 2$, $i = 1$, $j = 0$, one can improve (B.6), namely, one can show

$$\text{(B.8)} \qquad n^{-\kappa} \sup_{t\in[0,T]} U_{\lfloor nt\rfloor} \xrightarrow{\mathbb{P}} 0 \qquad \text{as } n \to \infty \text{ for } \kappa > 1,$$

see Barczy et al. [4].

C A version of the continuous mapping theorem

A function $f : \mathbb{R}_+ \to \mathbb{R}^d$ is called *càdlàg* if it is right continuous with left limits. Let $\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$ and $\mathsf{C}(\mathbb{R}_+, \mathbb{R}^d)$ denote the space of all $\mathbb{R}^d$-valued càdlàg and continuous functions on $\mathbb{R}_+$, respectively. Let $\mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d))$ denote the Borel σ-algebra on $\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$ for the metric

defined in Jacod and Shiryaev [8, Chapter VI, (1.26)] (with this metric $\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$ is a complete and separable metric space and the topology induced by this metric is the so-called Skorokhod topology). For $\mathbb{R}^d$-valued stochastic processes $(\boldsymbol{\mathcal{Y}}_t)_{t\in\mathbb{R}_+}$ and $(\boldsymbol{\mathcal{Y}}_t^{(n)})_{t\in\mathbb{R}_+}$, $n \in \mathbb{N}$, with càdlàg paths we write $\boldsymbol{\mathcal{Y}}^{(n)} \stackrel{\mathcal{D}}{\longrightarrow} \boldsymbol{\mathcal{Y}}$ if the distribution of $\boldsymbol{\mathcal{Y}}^{(n)}$ on the space $(\mathsf{D}(\mathbb{R}_+, \mathbb{R}), \mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)))$ converges weakly to the distribution of $\boldsymbol{\mathcal{Y}}$ on the space $(\mathsf{D}(\mathbb{R}_+, \mathbb{R}), \mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)))$ as $n \to \infty$. Concerning the notation $\stackrel{\mathcal{D}}{\longrightarrow}$ we note that if ξ and ξ_n, $n \in \mathbb{N}$, are random elements with values in a metric space (E, d), then we also denote by $\xi_n \stackrel{\mathcal{D}}{\longrightarrow} \xi$ the weak convergence of the distributions of ξ_n on the space $(E, \mathcal{B}(E))$ towards the distribution of ξ on the space $(E, \mathcal{B}(E))$ as $n \to \infty$, where $\mathcal{B}(E)$ denotes the Borel σ-algebra on E induced by the given metric d.

The following version of continuous mapping theorem can be found for example in Kallenberg [9, Theorem 3.27].

C.1 Lemma. *Let (S, d_S) and (T, d_T) be metric spaces and $(\xi_n)_{n\in\mathbb{N}}$, ξ be random elements with values in S such that $\xi_n \stackrel{\mathcal{D}}{\longrightarrow} \xi$ as $n \to \infty$. Let $f : S \to T$ and $f_n : S \to T$, $n \in \mathbb{N}$, be measurable mappings and $C \in \mathcal{B}(S)$ such that $\mathbb{P}(\xi \in C) = 1$ and $\lim_{n\to\infty} d_T(f_n(s_n), f(s)) = 0$ if $\lim_{n\to\infty} d_S(s_n, s) = 0$ and $s \in C$. Then $f_n(\xi_n) \stackrel{\mathcal{D}}{\longrightarrow} f(\xi)$ as $n \to \infty$.*

For the case $S = \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$ and $T = \mathbb{R}^q$ (or $T = \mathsf{D}(\mathbb{R}_+, \mathbb{R}^q)$), where d, $q \in \mathbb{N}$, we formulate a consequence of Lemma C.1.

For functions f and f_n, $n \in \mathbb{N}$, in $\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$, we write $f_n \stackrel{\mathrm{lu}}{\longrightarrow} f$ if $(f_n)_{n\in\mathbb{N}}$ converges to f locally uniformly, i.e., if $\sup_{t\in[0,T]} \|f_n(t) - f(t)\| \to 0$ as $n \to \infty$ for all $T > 0$. For measurable mappings $\Phi : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^q$ (or $\Phi : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathsf{D}(\mathbb{R}_+, \mathbb{R}^q)$) and $\Phi_n : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^q$ (or $\Phi_n : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathsf{D}(\mathbb{R}_+, \mathbb{R}^q)$), $n \in \mathbb{N}$, we will denote by $C_{\Phi,(\Phi_n)_{n\in\mathbb{N}}}$ the set of all functions $f \in \mathsf{C}(\mathbb{R}_+, \mathbb{R}^d)$ such that $\Phi_n(f_n) \to \Phi(f)$ (or $\Phi_n(f_n) \to\stackrel{\mathrm{lu}}{\longrightarrow} \Phi(f)$) whenever $f_n \stackrel{\mathrm{lu}}{\longrightarrow} f$ with $f_n \in \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$, $n \in \mathbb{N}$.

We will use the following version of the continuous mapping theorem several times, see, e.g., Barczy et al. [2, Lemma 4.2] and Ispány and Pap [6, Lemma 3.1].

C.2 Lemma. *Let $d, q \in \mathbb{N}$, and $(\boldsymbol{\mathcal{U}}_t)_{t\in\mathbb{R}_+}$ and $(\boldsymbol{\mathcal{U}}_t^{(n)})_{t\in\mathbb{R}_+}$, $n \in \mathbb{N}$, be $\mathbb{R}^d$-valued stochastic processes with càdlàg paths such that $\boldsymbol{\mathcal{U}}^{(n)} \stackrel{\mathcal{D}}{\longrightarrow} \boldsymbol{\mathcal{U}}$. Let $\Phi : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^q$ (or $\Phi : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathsf{D}(\mathbb{R}_+, \mathbb{R}^q)$) and $\Phi_n : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^q$ (or $\Phi_n : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathsf{D}(\mathbb{R}_+, \mathbb{R}^q)$), $n \in \mathbb{N}$, be measurable mappings such that there exists $C \subset C_{\Phi,(\Phi_n)_{n\in\mathbb{N}}}$ with $C \in \mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d))$ and $\mathbb{P}(\boldsymbol{\mathcal{U}} \in C) = 1$. Then $\Phi_n(\boldsymbol{\mathcal{U}}^{(n)}) \stackrel{\mathcal{D}}{\longrightarrow} \Phi(\boldsymbol{\mathcal{U}})$.*

In order to apply Lemma C.2, we will use the following statement several times, see Barczy et al. [4, Lemma B.3].

C.3 Lemma. *Let $d, p, q \in \mathbb{N}$, $h : \mathbb{R}^d \to \mathbb{R}^q$ be a continuous function and $K : [0, 1] \times \mathbb{R}^{2d} \to \mathbb{R}^p$ be a function such that for all $R > 0$ there exists $C_R > 0$ such that*

$$\|K(s, x) - K(t, y)\| \leqslant C_R \left(|t - s| + \|x - y\|\right) \tag{C.1}$$

for all $s, t \in [0,1]$ *and* $x, y \in \mathbb{R}^{2d}$ *with* $\|x\| \leqslant R$ *and* $\|y\| \leqslant R$. *Moreover, let us define the mappings* $\Phi, \Phi_n : \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^{q+p}$, $n \in \mathbb{N}$, *by*

$$\Phi_n(f) := \left(h(f(1)), \frac{1}{n} \sum_{k=1}^{n} K\left(\frac{k}{n}, f\left(\frac{k}{n}\right), f\left(\frac{k-1}{n}\right)\right)\right),$$

$$\Phi(f) := \left(h(f(1)), \int_0^1 K(u, f(u), f(u))\,\mathrm{d}u \right)$$

for all $f \in \mathsf{D}(\mathbb{R}_+, \mathbb{R}^d)$. *Then the mappings* Φ *and* Φ_n, $n \in \mathbb{N}$, *are measurable, and* $C_{\Phi,(\Phi_n)_{n\in\mathbb{N}}} = \mathsf{C}(\mathbb{R}_+, \mathbb{R}^d) \in \mathcal{B}(\mathsf{D}(\mathbb{R}_+, \mathbb{R}^d))$.

D Convergence of random step processes

We recall a result about convergence of random step processes towards a diffusion process, see Ispány and Pap [6]. This result is used for the proof of convergence (5.1).

D.1 Theorem. *Let* $\boldsymbol{\gamma} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times r}$ *be a continuous function. Assume that uniqueness in the sense of probability law holds for the SDE*

$$\mathrm{d}\boldsymbol{\mathcal{U}}_t = \gamma(t, \boldsymbol{\mathcal{U}}_t)\,\mathrm{d}\boldsymbol{\mathcal{W}}_t, \qquad t \in \mathbb{R}_+, \tag{D.1}$$

with initial value $\boldsymbol{\mathcal{U}}_0 = \boldsymbol{u}_0$ *for all* $\boldsymbol{u}_0 \in \mathbb{R}^d$, *where* $(\boldsymbol{\mathcal{W}}_t)_{t\in\mathbb{R}_+}$ *is an* r*-dimensional standard Wiener process. Let* $(\boldsymbol{\mathcal{U}}_t)_{t\in\mathbb{R}_+}$ *be a solution of* (D.1) *with initial value* $\boldsymbol{\mathcal{U}}_0 = \boldsymbol{0} \in \mathbb{R}^d$.

For each $n \in \mathbb{N}$, *let* $(\boldsymbol{U}_k^{(n)})_{k\in\mathbb{N}}$ *be a sequence of* d*-dimensional martingale differences with respect to a filtration* $(\mathcal{F}_k^{(n)})_{k\in\mathbb{Z}_+}$, *i.e.,* $\mathbb{E}(\boldsymbol{U}_k^{(n)} \mid \mathcal{F}_{k-1}^{(n)}) = 0$, $n \in \mathbb{N}$, $k \in \mathbb{N}$. *Let*

$$\boldsymbol{\mathcal{U}}_t^{(n)} := \sum_{k=1}^{\lfloor nt \rfloor} \boldsymbol{U}_k^{(n)}, \qquad t \in \mathbb{R}_+, \quad n \in \mathbb{N}.$$

Suppose that $\mathbb{E}\left(\|\boldsymbol{U}_k^{(n)}\|^2\right) < \infty$ *for all* $n, k \in \mathbb{N}$. *Suppose that for each* $T > 0$,

(i) $\sup\limits_{t\in[0,T]} \left\| \sum\limits_{k=1}^{\lfloor nt \rfloor} \mathbb{E}\Big(\boldsymbol{U}_k^{(n)} (\boldsymbol{U}_k^{(n)})^\top \mid \mathcal{F}_{k-1}^{(n)}\Big) - \int_0^t \boldsymbol{\gamma}(s, \boldsymbol{\mathcal{U}}_s^{(n)}) \boldsymbol{\gamma}(s, \boldsymbol{\mathcal{U}}_s^{(n)})^\top \mathrm{d}s \right\| \stackrel{\mathbb{P}}{\longrightarrow} 0,$

(ii) $\sum\limits_{k=1}^{\lfloor nT \rfloor} \mathbb{E}\left(\|\boldsymbol{U}_k^{(n)}\|^2 \mathbb{1}_{\{\|\boldsymbol{U}_k^{(n)}\| > \theta\}} \,\middle|\, \mathcal{F}_{k-1}^{(n)}\right) \stackrel{\mathbb{P}}{\longrightarrow} 0$ *for all* $\theta > 0$,

where $\stackrel{\mathbb{P}}{\longrightarrow}$ *denotes convergence in probability. Then* $\boldsymbol{\mathcal{U}}^{(n)} \stackrel{\mathcal{D}}{\longrightarrow} \boldsymbol{\mathcal{U}}$ *as* $n \to \infty$.

Note that in (i) of Theorem D.1, $\|\cdot\|$ denotes a matrix norm, while in (ii) it denotes a vector norm.

References

[1] Athreya, K. B. and Ney, P. E. (1972). *Branching Processes*, Springer-Verlag, New York-Heidelberg.

[2] Barczy, M., Ispány, M. and Pap, G. (2010). Asymptotic behavior of CLS estimators of autoregressive parameter for nonprimitive unstable INAR(2) models. Available on the ArXiv: `http://arxiv.org/abs/1006.4641`.

[3] Barczy, M., Ispány, M. and Pap, G. (2011). Asymptotic behavior of unstable INAR(p) processes. *Stochastic Process. Appl.* **121(3)** 583–608.

[4] Barczy, M., Ispány, M. and Pap, G. (2012). Asymptotic behavior of CLS estimators for unstable INAR(2) models.
Available on the ArXiv: `http://arxiv.org/abs/1202.1617`.

[5] Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press, Princeton.

[6] Ispány, M. and Pap, G. (2010). A note on weak convergence of step processes. *Acta Mathematica Hungarica* **126(4)** 381–395.

[7] Ispány, M. and Pap, G. (2012). Asymptotic behavior of critical primitive multi-type branching processes with immigration. Available on the ArXiv: `http://arxiv.org/abs/1205.0388`.

[8] Jacod, J. and Shiryaev, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. Springer-Verlag, Berlin.

[9] Kallenberg, O. (1997). *Foundations of Modern Probability*. Springer, New York, Berlin, Heidelberg.

[10] Kesten, H. and Stigum, B. P. (1966). A limit theorem for multidimensional Galton–Watson processes. *Ann. Math. Statist.* **37(5)** 1211–1223.

[11] Musiela, M. and Rutkowski, M. (1997). *Martingale Methods in Financial Modelling*, Springer-Verlag, Berlin, Heidelberg.

[12] Putzer, E. J. (1966). Avoiding the Jordan canonical form in the discussion of linear systems with constant coefficients. *The American Mathematical Monthly* **73(1)** 2–7.

[13] Quine, M. P. (1970). The multi-type Galton–Watson process with immigration. *J. Appl. Probab.* **7(2)** 411–422.

[14] Revuz, D. and Yor, M. (2001). *Continuous Martingales and Brownian Motion*, 3rd ed., corrected 2nd printing. Springer-Verlag, Berlin.

[15] Tanaka, K. (1996). *Time Series Analysis, Nonstationary and Noninvertible Distribution Theory*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., New York.

[16] WEI, C. Z. and WINNICKI, J. (1989). Some asymptotic results for the branching process with immigration- *Stochastic Process. Appl.* **31(2)** 261–282.

[17] WEI, C. Z. and WINNICKI, J. (1990). Estimation of the means in the branching process with immigration. *Ann. Statist.* **18** 1757–1773.

[18] WINNICKI, J. (1991). Estimation of the variances in the branching process with immigration. *Probability Theory and Related Fields* **88(1)** 77–106.